

Ein Bewertungsrahmen für fundierte Sprachdiagnostik in der Kita

Qualitätsmerkmale

für Sprachstandsverfahren im Elementarbereich

Michael Becker-Mrotzek
Konrad Ehlich
Iris Füssenich
Hartmut Günther
Marcus Hasselhorn
Michaela Hopf
Stefan Jeuk
Drorit Lengyel
Uwe Neugebauer
Argyro Panagiotopoulou
Petra Stanat
Jürgen Wilbert

Herausgegeben vom

Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache



6	DAS WICHTIGSTE IN KÜRZE
9	RELEVANZ
14	VORGEHEN
16	WAS QUALITÄTSMERKMALE LEISTEN KÖNNEN
20	QUALITÄTSMERKMALE
20	Berücksichtigung sprachlicher Basisqualifikationen
23	Validität
24	Objektivität
26	Reliabilität
27	Normierung
28	Fehlerquote (Gewichtung der Auswahlfehler)
30	Qualifizierung der pädagogischen Fachkräfte
32	Zeitliche Anforderungen – Durchführung, Auswertung & Ergebnismeldung
33	Mehrsprachigkeit
35	Spezifität der Diagnostik
37	HANDLUNGSEMPFEHLUNGEN
40	AUTORINNEN UND AUTOREN
41	LITERATUR
42	IMPRESSUM



Prof. Barbara John

Ehemalige Ausländerbeauftragte des Berliner Senats (1981–2003) und Mitglied im wissenschaftlichen Beirat des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache

VORWORT

Wir alle blieben lebenslang sprachlos, kämen wir nicht zur Welt mit einer angeborenen Neugier auf Sprache. Deshalb lernen die meisten Kinder ganz natürlich sprechen: Auf dem Spielplatz erweitern sie ihren Wortschatz, beim Mittagessen in der Kita verbessern sie den Satzbau und beim Vorlesen und Erzählen erkennen sie, wie eine Geschichte aufgebaut ist. Doch was passiert, wenn sie sich die Welt zu wenig sprachlich erschließen, weil Anregungen fehlen? Sie bleiben unter ihren intellektuellen Fähigkeiten – manchmal ein Leben lang. So können am Ende der zehnten Klasse bis zu 20 Prozent der Schülerinnen und Schüler nur unzureichend lesen, wie die PISA-Studien zeigen. Das muss nicht so sein, würden diese Kinder, ob mit oder ohne Einwanderungsbiografie, schon in der frühkindlichen Entwicklung sprachlich gefördert.

Förderprogramme gibt es wie Sand am Meer. Aber wir wissen längst, dass es nicht reicht, die Kleinsten ungezielt drauflos sprachlich zu fördern, ohne ihren speziellen Bedarf zu kennen. Zwar sind bundesweit in den Kitas zahlreiche Diagnoseinstrumente im Einsatz, die Aufschluss über die sprachliche Entwicklung der Kinder geben sollen. Doch ob diese sogenannten Sprachstandsverfahren das wirklich leisten, darf nach vielen Studien bezweifelt werden. Es ist enttäuschend, aber nicht zu bestreiten: Bisher gibt es keine verbindlichen wissenschaftlichen Standards für die Sprachdiagnostik im Elementarbereich.

Spracherwerb ist bekanntermaßen ein äußerst komplexer Prozess. Noch wissen wir wenig darüber. Es hilft der Praxis aber nicht weiter, den Prozess sprachlicher Entwicklung in seiner Ganzheit zunächst entschlüsseln zu wollen und dann erst zu handeln. Schon jetzt braucht es wissenschaftlich fundierte Lösungen für konkrete Probleme. Ein Problem, das lösbar erscheint, ist die genaue Auswahl der Kinder, die einen Sprachförderbedarf haben. Ich freue mich daher, dass auf Initiative des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache Wissenschaftlerinnen und Wissenschaftler aus zahlreichen Disziplinen gemeinsam Qualitätsmerkmale für Sprachstandsverfahren im Elementarbereich entwickelt haben. Ich wünsche mir, dass diese Kriterien in Politik und in Wissenschaft breit diskutiert werden, denn sie können Ausgangspunkt sein für nötige strukturelle Veränderungen – in der Weiterentwicklung der Sprachdiagnostik, in der Aus- und Fortbildung der Erzieherinnen und Erzieher und bei der Förderpraxis.

Damit kann das Mercator-Institut den politisch Verantwortlichen ein kompetenter Begleiter auf dem Weg zu einer besseren sprachlichen Bildung sein.

Barbara John

DAS WICHTIGSTE IN KÜRZE

HINTERGRUND

Eine bedarfsgerechte Sprachförderung muss möglichst früh ansetzen, um allen Kindern optimale Startchancen in ihre Schullaufbahn zu ermöglichen. Dass das insbesondere die Sprachdiagnostik im Elementarbereich betrifft, hat die Bildungspolitik bereits erkannt: Zwischen 2002 und 2008 haben 14 Bundesländer Verfahren eingeführt, um den sprachlichen Entwicklungsstand der vier- bis fünfjährigen Kinder festzustellen und angemessene Sprachfördermaßnahmen noch vor dem Schuleintritt einzuleiten.

Die Qualität und Wirksamkeit dieser Verfahren sind jedoch fraglich: Bundesweit sind 17 verschiedene Sprachstandsverfahren im Einsatz, die Förderquoten schwanken je nach Bundesland zwischen zehn und 50 Prozent. Ein Beispiel: In neun Bundesländern ist die Sprachdiagnostik vor Schuleintritt verpflichtend, in fünf Ländern wird jedoch nur der Sprachstand von bestimmten Gruppen, z. B. von Kindern mit Migrationshintergrund, ermittelt. Eine empirische Überprüfung der Verfahren ist bisher nur vereinzelt erfolgt, die Vergleichbarkeit wurde bislang nicht übergreifend überprüft. So kommt auch der Bildungsbericht 2012 zu dem Schluss, dass „die Länder nach wie vor auffallend unterschiedlich vorgehen, was zu einem uneinheitlichen Umgang mit den Kindern führt.“

ZIEL UND VORGEHEN DER UNTERSUCHUNG

Bund und Länder haben mehrfach angekündigt, einheitliche Standards für Sprachstandsverfahren zu entwickeln, zuletzt im Nationalen Aktionsplan Integration von 2012. Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache unterstützt dieses Ziel. Daher hat es eine interdisziplinäre Expertenkommission initiiert, die erstmals wissenschaftliche Qualitätsmerkmale für Sprachstandsverfahren im Elementarbereich entwickelt hat. Dieser Bewertungsrahmen definiert aus psychologischer, pädagogischer, sprachtherapeutischer und linguistischer Perspektive, welche Kriterien für Sprachstandsverfahren gelten sollten, um den Sprachstand jedes Kindes sowohl vergleichbar als auch wissenschaftlich fundiert zu ermitteln.

Sie umfassen

die sprachlichen Schlüsselkompetenzen, die für einen guten Start in die Schullaufbahn und die Teilhabe an der Gesellschaft von zentraler Bedeutung sind;

fundierte Hinweise darauf, wie Sprachstandsverfahren Kindern gerecht werden können, die Deutsch als Zweitsprache erlernen, ohne Mehrsprachigkeit zu stigmatisieren;

empirische Gütemerkmale, die für eine kontinuierliche Evaluation und Optimierung der Sprachstandsverfahren unabdingbar sind;

die Qualifizierung der pädagogischen Fachkräfte, die für die Sprachdiagnostik zuständig sind;

Orientierungswerte, wie sich die wissenschaftlich-pädagogischen Anforderungen an eine wirkungsvolle Diagnostik mit dem pädagogischen Alltag in Einklang bringen lassen.

Zur Expertenkommission gehören Prof. Dr. Michael Becker-Mrotzek (Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache), Prof. Dr. Dr. h. c. Konrad Ehlich (Freie Universität Berlin), Prof. Dr. Iris Füssenich (Pädagogische Hochschule Ludwigsburg), Prof. Dr. Hartmut Günther (Universität zu Köln), Prof. Dr. Marcus Hasselhorn (Deutsches Institut für Internationale Pädagogische Forschung (DIPF)), Dr. Michaela Hopf (Deutsches Jugendinstitut (DJI)), apl. Prof. Dr. Stefan Jeuk (Pädagogische Hochschule Ludwigsburg), Prof. Dr. Drorit Lengyel (Universität Hamburg), Dr. Uwe Neugebauer (Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache), Prof. Dr. Argyro Panagiotopoulou (Universität zu Köln), Prof. Dr. Petra Stanat (Institut zur Qualitätsentwicklung im Bildungswesen (IQB)) und Dr. habil. Jürgen Wilbert (Universität zu Köln).

DIE QUALITÄTSMERKMALE

Die Expertenkommission hat für zehn Handlungsfelder Qualitätsmerkmale entwickelt. Für jedes Merkmal ist eindeutig definiert, ab wann ein Sprachstandsverfahren das Merkmal ausreichend, gut oder sehr gut erfüllt. Für das Handlungsfeld Mehrsprachigkeit beispielsweise ist ein Verfahren dann ausreichend, wenn u. a. die Erstsprache des Kindes erfasst wird, die Kontaktmonate mit

der Zweitsprache und die Einschätzung der Eltern, wie das Kind Erst- und Zweitsprache spricht bzw. kombiniert.

Berücksichtigung sprachlicher Basisqualifikationen

Das Verfahren erfasst die zentralen Bereiche der sprachlichen Entwicklung eines Kindes.

Validität

Das Verfahren ist so konstruiert, dass es exakt den Sprachstand eines Kindes erfasst und nicht andere Bereiche, wie zum Beispiel kulturelles Wissen.

Objektivität

Das Verfahren ist unabhängig von subjektiven Einflussfaktoren, wie beispielsweise dem Verhalten der pädagogischen Fachkraft.

Reliabilität

Das Verfahren misst den Sprachstand des Kindes präzise. Wird das Verfahren wiederholt, steht am Ende das gleiche Ergebnis.

Normierung

Die Ergebnisse eines Kindes lassen sich mit den Ergebnissen anderer Kinder vergleichen. Nur so lässt ein Verfahren eine Aussage darüber zu, welche Kinder einen Sprachförderbedarf haben und welche Kinder sich sprachlich altersgerecht entwickeln.

Fehlerquote (Gewichtung der Auswahlfehler)

Das Verfahren stellt sicher, dass die Kinder, die einen Sprachförderbedarf haben, auch wirklich identifiziert werden. Das Verfahren ist besonders sensitiv, d. h. im Zweifelsfall werden zunächst mehr Kinder mit einem möglichen Sprachförderbedarf identifiziert, der sich später in der detaillierten Diagnose des Förderbedarfs ggf. nicht mehr bestätigt.

Qualifizierung der pädagogischen Fachkräfte

Die Personen, die den Sprachstand der Kinder messen, sind im Vorfeld für Sprachdiagnostik sensibilisiert und qualifiziert worden.

Zeitliche Anforderungen

Das Verfahren berücksichtigt die Konzentrationsfähigkeit des Kindes und die zeitliche Machbarkeit im pädagogischen Alltag.

Mehrsprachigkeit

Das Verfahren erfasst die besonderen Rahmenbedingungen von Kindern, die Deutsch als Zweitsprache erlernen. Nicht jedes mehrsprachige Kind hat per se einen Sprachförderbedarf.

Spezifität der Diagnostik

Das Verfahren gibt Hinweise darauf, welche Sprachfördermaßnahmen für das Kind in Frage kommen.

DIE EMPFEHLUNGEN

Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache lädt die verantwortlichen Akteure ein, ihre Verfahren kritisch zu prüfen. Die folgenden Handlungsempfehlungen bieten Impulse für einen Dialog zwischen Wissenschaft, Bildungspolitik und Praxis.

Sprachstandsverfahren sollten die zentralen Bereiche der sprachlichen Entwicklung eines Kindes sowie seiner (mehrsprachigen) Sprachbiografie erfassen. Gleichzeitig müssen sie wissenschaftliche Qualitätsmerkmale berücksichtigen und im pädagogischen Alltag praktikabel sein. Um diesen vielfachen Anforderungen gerecht zu werden, **empfiehlt das Mercator-Institut, ein interdisziplinäres Gremium ins Leben zu rufen.** Neben Vertreterinnen und Vertretern aus den betreffenden wissenschaftlichen Disziplinen sollten ebenso Akteure aus der Testentwicklung, der pädagogischen Praxis und der Bildungsadministration beteiligt sein.

Die stark variierenden Förderquoten weisen darauf hin, dass die Bundesländer bei der Definition eines Sprachförderbedarfs stark voneinander abweichen. Damit die

Chance auf Förderung nicht vom Bundesland abhängt, empfiehlt das Mercator-Institut eine zentrale Koordinierungsstelle zur einheitlichen Definition, ab wann Sprachförderung nötig ist. Ein solches Gremium kann beispielsweise bei den Fachministerkonferenzen der Länder oder in der Bund-Länder-Initiative „Bildung durch Sprache und Schrift“ angesiedelt sein und auf Basis der Empfehlungen des Expertengremiums Hinweise zur Qualitätsverbesserung der landesspezifisch eingesetzten Verfahren geben.

Pädagogische Fachkräfte müssen nicht nur für die Durchführung der Sprachstandsverfahren geschult werden, sondern benötigen für eine fundierte Diagnostik ebenso allgemeine Kenntnisse über den kindlichen Spracherwerb. Sie sollten daher **gezielt auf die Durchführung der Sprachstandsverfahren vorbereitet werden.** Sowohl in der beruflichen wie auch akademischen Ausbildung braucht es Module, die angehende Fachkräfte, entsprechend ihres Profils, auf Sprachdiagnostik und Sprachförderung vorbereiten.

Die in den Ländern verpflichtend eingeführten Verfahren sollten einem **Abgleich mit den hier vorgestellten Qualitätsmerkmalen unterzogen und darauf basierend weiterentwickelt werden.** In einer Querschnittsstudie müssten dafür zunächst die einzelnen Verfahren unter gleichen Einsatz- und Rahmenbedingungen auf Basis der hier erarbeiteten Qualitätsmerkmale überprüft werden. Mithilfe einer **Längsschnittstudie** ließe sich dann überprüfen, ob die Kinder, die im Alter von vier Jahren ggf. eine Sprachförderung erhalten haben, am Ende des ersten Schuljahres über die für dieses Alter definierten sprachlichen Kompetenzen verfügen.

RELEVANZ

Seit Ende der 1990er-Jahre wurde im Zusammenhang mit Diskussionen über die Bildungsbenachteiligung von Kindern mit Migrationshintergrund¹ auch zunehmend der Elementarbereich in die Sprachdiagnostik und -förderung eingebunden (vgl. Reich 2006, S. 915), wobei Sprachdiagnosen im Elementarbereich bereits seit Langem im Rahmen der Sprachtherapie etabliert sind. Die Legitimation hierfür ist leicht nachzuvollziehen: **Wird der (Zweit-)Spracherwerb frühzeitig nachhaltig unterstützt und werden adäquate Fördermaßnahmen initiiert, können allen Kindern optimale Startchancen zu Beginn der Primarstufe ermöglicht werden** (Kniffka/Siebert-Ott 2007, S. 122).

Im Nationalen Integrationsplan von 2007 verpflichten sich die Bundesländer, vor der Einschulung Verfahren zur Feststellung und Beobachtung des Sprachstandes² einzusetzen und je nach Bedarf eine Sprachförderung zu ermöglichen (Die Bundesregierung 2007, S. 52 ff.). Da die Bildungshoheit bei den einzelnen Bundesländern liegt, empfehlen diese unterschiedliche Verfahren zur Sprachstandsdiagnostik und Sprachförderung in Kindertagesstätten, so dass sich eine „[...] große Heterogenität zwischen den verschiedenen Ländern bei der Umsetzung dieser Vorhaben“ (Lisker 2010, S. 3) zeigt: Zum einen stellt sich die Frage, ob Sprachstandserhebungen verpflichtend für alle Kinder durchgeführt werden sollen und in welchem Alter jene veranlasst werden. Zum anderen muss entschieden werden, welche Verfahren eingesetzt werden und welche Konsequenzen aus den Ergebnissen entstehen (vgl. ebd.).

In den Jahren 2002 bis 2008 wurden in nahezu allen Bundesländern Sprachstandsverfahren für die Altersgruppe der vier- bis fünfjährigen Kinder vereinbart, **in neun Bundesländern sind jene verpflichtend**. Allerdings müssen in fünf Bundesländern nur bestimmte Personengruppen daran teilnehmen, so z. B. in Bayern und Hessen „alle Kinder mit Migrationshintergrund“ (Bildungsbericht 2012, S. 248). Hier manifestieren sich Gleichsetzungen, die der Realität nicht gerecht werden: „Warum fragt man nach der Herkunft, wenn man die Sprache meint?“ (vgl. Chlosta und Ostermann 2007) Ebenso schwankt die Quote von als förderbedürftig klassifizierten Kindern zwischen zehn Prozent (Saarland 2008) und etwas über 50 Prozent (Bremerhaven 2008 und 2009). Zu einzelnen Sprachstandsverfahren existieren empirisch angelegte Untersuchungen. So hat Settinieri³ (2012) die Ergebnisse von SSMIK (Bayern) und Delfin 4 (Nordrhein-Westfalen) an einer Stichprobe von 332 Kindern verglichen und kommt zu dem Ergebnis, dass sie knapp zur Hälfte (47,4 Prozent) das Gleiche messen: „Anders gesagt bedeutet das, dass die beiden Verfahren zum größeren Teil etwas Unterschiedliches, also entweder nicht den Sprachstand der Kinder oder zumindest völlig andere Aspekte des Konstrukts messen“ (Settinieri 2012, S. 44).

ÜBERSICHT ÜBER SPRACHSTANDSVERFAHREN IN DEN BUNDESLÄNDERN

Sprachstandsverfahren	Bundesland
Cito	Bremen
Delfin 4	Nordrhein-Westfalen, Sachsen-Anhalt
Deutsch Plus 4	Berlin
Fit in Deutsch	Niedersachsen
Früh Deutsch lernen	Saarland
HASE	Baden-Württemberg
HAVAS 5	Schleswig-Holstein
Kenntnisse in DaZ erfassen	Bayern
KiSS	Hessen
KISTE	Brandenburg
Meilensteine der Sprachentwicklung	Brandenburg
QuaSta	Berlin
SETK 3-5	Baden-Württemberg
SSV	Sachsen
SISMIK	Bayern, Schleswig-Holstein (unterschiedliche Angaben)
VER-ES	Rheinland-Pfalz
Vorstellungsverfahren Viereinhalbjähriger Bildimpulse	Hamburg

Tabelle 1

Hinweis: Vier Bundesländer haben zwei bzw. drei (BW) Verfahren zugelassen; zwei Bundesländer empfehlen dasselbe Verfahren (Delfin 4); BB hat 2013 von WESPE auf Meilensteine der Sprachentwicklung umgestellt.

Quellen: (a) Lisker, Andrea (2010): Sprachstandsfeststellung und Sprachförderung im Kindergarten sowie beim Übergang in die Schule. Expertise im Auftrag des Deutschen Jugendinstituts; (b) Bildungsbericht 2012, S. 248

Die Qualität eines diagnostischen Verfahrens ist für die pädagogischen Fachkräfte wie auch für die Eltern nicht immer leicht ersichtlich, weil einerseits häufig nur ausgewählte Merkmale der Verfahren bekanntgegeben werden und die Funktionsweise der Instrumente nicht ausreichend dargestellt wird. Andererseits sind die pädagogischen Fachkräfte, die die Verfahren durchführen, häufig nicht ausreichend geschult. Damit ist es für die

Beteiligten schwierig, professionell entwickelte Instrumente von solchen zu unterscheiden, die wissenschaftliche Qualitätsmerkmale nur teilweise erfüllen.

Nationaler und internationaler Forschungsstand

In der Psychologie ist die Entwicklung von Testverfahren ein Forschungsfeld, für das es einen empirisch relativ gut gesicherten Kenntnisstand gibt. Gleichzeitig exis-

tiert in der Linguistik keine einheitliche Spracherwerbstheorie, die als Grundlage für Sprachstandsverfahren dienen könnte. Welche sprachlichen Kompetenzen ein Verfahren berücksichtigen muss, unterliegt also einer gewissen Vorläufigkeit – hier bedarf es weiterer Forschung.

Die Qualität psychodiagnostischer Verfahren wurde sowohl international⁴ als auch national⁵ betrachtet. Ein solches System der Testbewertung stellt insbesondere in Deutschland eine Herausforderung dar, die in der Vergangenheit nur unzureichend bewältigt werden konnte (Kersting 2006). Neben den klassischen Gütekriterien werden dabei z. B. im Testbeurteilungssystem des Testkuratoriums (TBS-TK) der Föderation Deutscher Psychologinnenvereinigungen (2009) eine mehr oder weniger große Anzahl weiterer Kriterien berücksichtigt, so z. B. die theoretische Grundlage als Ausgangspunkt der Testkonstruktion wie auch die Zugänglichkeit von Informationen und deren Informationsgehalt, insbesondere bzgl. empirischer Ergebnisse zum Testverfahren. Das TBS-TK bewertet die einzelnen Bereiche dabei mit den vier Abstufungen „Der Test erfüllt die Anforderungen...“ voll, weitgehend, teilweise oder nicht.

Bewertungsraster für Sprachstandsverfahren im Elementarbereich sind im deutschsprachigen Raum z. B. von Kany und Schöler (2010) oder auch Kiese-Himmel und Rosenfeld (2012) entwickelt und eingesetzt worden, im US-amerikanischen Raum liegen Ansätze z. B. von Plante und Vance (1995) wie auch Law et al. (2000) vor. Während Kany und Schöler (2010) die Verfahren hinsichtlich ihrer Eignung für Kinder mit Spracherwerbsstörungen oder mehrsprachige Kinder betrachten, umfasst das Bewertungsraster von Kiese-Himmel und Rosenfeld (2012) insgesamt 15 Kategorien zur Bewertung der Qualität eines Verfahrens. Es enthält allerdings keine Angaben, ab wann ein Merkmal als erfüllt bewertet wird. Die 15 Kategorien waren u. a. Angaben zur Objektivität⁶, zu den unterschiedlichen Formen der Reliabilität, Validität⁷, zur Durchführungsdauer sowie den berücksichtigten sprachlichen Dimensionen.

Mangelnder Transfer in die Praxis

Etabliert haben sich diese Anforderungen an Sprachstandsverfahren im deutschsprachigen Raum offensichtlich weder hinsichtlich der Beachtung unterschiedlicher Zielgruppen noch hinsichtlich der aufgestellten Qualitätsmerkmale. Dies wird auch in dem vom Bundesministerium für Bildung und Forschung (BMBF) beauftragten Bericht „Anforderungen an Verfahren der regelmäßigen Sprachstandsfeststellung als Grundlage für die frühe und individuelle Förderung von Kindern, mit

und ohne Migrationshintergrund“ (Ehlich et al., 2007) deutlich – die bereits dort ausführlich und detailliert formulierten Anforderungen finden in den seitdem entwickelten Verfahren keine Berücksichtigung, z. B. in Bezug auf die Frage, wie Kinder Laute, Silben und Wörter wahrnehmen, unterscheiden und selbst nutzen (phonische Basisqualifikation). Dabei ist der Bereich der phonischen Qualifikation verhältnismäßig einfach erfassbar. Deutlich schwieriger stellt es sich dar, wenn z. B. die Erzählfähigkeit oder die Wortschatzentwicklung betrachtet wird (vgl. Tabelle 2, S. 10).

Die bisherigen Untersuchungen und Evaluierungen zahlreicher Sprachstandsverfahren (vgl. Ehlich et al. 2005; Fried 2005; Lisker 2010) haben auf erhebliche Mängel bzw. Unklarheiten bei deren Entwicklung und Einsatz hingewiesen. Eine zentrale Schwierigkeit besteht darin, dass die Forschung bisher zu wenig über die kindliche Sprachaneignung, insbesondere in Bezug auf die sprachlichen Basisqualifikationen und deren Zusammenhänge, weiß. Darüber hinaus liegt ein von Wissenschaft und Praxis getragener Konsens, wie der natürliche Spracherwerb eines Kindes verläuft und ab wann ein Kind von dieser normalen Entwicklung abweicht, bisher nicht bzw. allenfalls in Ansätzen vor.

Gleichwohl sind solche Normalitätserwartungen eine gesellschaftliche Wirklichkeit: Der Sprachstand der Kinder wird anhand dieser Vorstellungen untersucht und eingeordnet. Deshalb ist es erforderlich, diese Normalitätserwartungen möglichst präzise zu erfassen und kritisch zu reflektieren. Sprachstandsverfahren sind daran zu überprüfen, ob und in welchem Umfang sie dies berücksichtigen. Der Blick in die Praxis zeigt: Kinder werden verpflichtet, an Sprachstandsverfahren teilzunehmen, die nicht ausreichend wissenschaftlichen Standards genügen und – je nach Bundesland – äußerst unterschiedlichen psychometrischen und sprachdidaktischen Vorstellungen folgen oder die Auswahl in nur sehr ungenügender Weise vornehmen.

VERANKERUNG SPRACHLICHER BASISQUALIFIKATIONEN IN AUSGEWÄHLTEN SPRACHSTANDSVERFAHREN				
Wie werden einzelne sprachliche Schlüsselkompetenzen untersucht?				
Sprachstandsverfahren	Merkmal I: „Phonologische Bewusstheit“ ⁸ als Facette der Phonischen Basisqualifikation	Bewertungskategorien	Merkmal II: „Wortschatzentwick- lung“ als Facette der Semantischen Basisqualifikation	Bewertungskategorien
QuaSta – Beobachtungs- bzw. Beurteilungsbogen	3 Beurteilungsbereiche (auditive Merkfähigkeit, Silbengliederung, Reimbildung): Auditive Merkfähigkeit: Wörter nachsprechen, Fantasiewörter nach- sprechen	Bewertung 4-fach abgestuft: meistens, häufig, selten, nie	Bereich „Wortschatzerweiter- ung“, 3 Bereiche anzukreuzen: „sich für neue Wörter interessieren“; „sich neue Wörter aneignen“ sowie „Oberbegriffe verstehen“	Bewertung 4-fach abgestuft: meistens, häufig, selten, nie
Früh Deutsch lernen – Beobachtungs- bzw. Beurteilungsbogen	„Sprachfehler? Falls ja, welcher?“	Bewertung 2-fach abge- stuft: ja / nein	3 Items: Das Kind kann ... a) Gegenstände des täglichen Lebens in der deutschen Sprache benennen. b) Mengen (bis 5) in deutscher Sprache benennen und in deutscher Sprache bis _____ zählen. c) Farben erkennen und in deutscher Sprache benennen.	Bewertung 2-fach abge- stuft: „ausreichend“ – „nicht ausreichend“
Delfin 4 (1. Stufe) – Screening	1 Aufgabentyp mit 8 Items pro Kind: Non-Wörter nachspre- chen	Pro korrekt nachgespro- chenem Non-Wort 1 Punkt	Nicht berücksichtigt in Stufe 1	–
Meilensteine der Sprachentwicklung – Beobachtungsbogen	Nicht berücksichtigt	–	1 Item (48 Monate): „Der Wortschatz des Kindes behindert nicht das Verstehen und Sprechen in alltäglichen Situationen und bei alltäglichen Handlun- gen. Wortfindungs- schwierigkeiten beim Benennen von alltäglichen Dingen und Objekten (Objektbegrif- fen) sind selten.“	Antwortoptionen: ja / nein

Tabelle 2

1 Im Folgenden wird der Begriff „mehrsprachige Kinder“ bzw. „mehrsprachig aufwachsende Kinder“ verwendet, der die Realität besser widerspiegelt. Nur in Fällen, in denen explizit auf diesen Begriff zurückgegriffen werden muss, wird er genutzt, so z. B. bei der Definition der Gruppen, bei denen das Sprachverfahren in einem Bundesland eingesetzt wird.

2 Der Sprachstand kann in jedem Verfahren natürlich nur in Ausschnitten betrachtet werden – der Sprachstand insgesamt kann nicht erfasst werden.

3 Ahrenholz, B., Knapp, W., Becker, T. und Jeuk, S. (Hrsg.) (2012). Sprachstand erheben – Spracherwerb erforschen.

4 Z. B. von der Internationalen Testkommission (ITC) oder vom Joint Committee der American Psychological Association, American Educational Research Association und des National Council on Measurement in Education (<http://www.apa.org/science/standards.html> sowie

<http://www.teststandards.org/>). Zwei weitere allgemeine Ansätze zur Qualitätseinschätzung sind durch das EFPA „Review Model for the description and evaluation of psychological Tests“ und die Expertengruppe um Bartram (2008) sowie das COTAN-System (Evers 2001) gegeben.

5 Z. B. vom Testkuratorium (<http://www.zpid.de/index.php?wahl=Testkuratorium>)

6 Standardisierung der Auswertung und Durchführung

7 U. a. prädiktive Validität, konvergente und diskriminante Validität, Klassifikations-Validität etc.

8 Unter phonologischer Bewusstheit versteht man die Fähigkeit, die Aufmerksamkeit auf die Laute der Sprache zu richten, und zwar unabhängig von ihrer Bedeutung, und diese Laute bewusst zu manipulieren. Erfasst wird diese Fähigkeit typischerweise mit Aufgaben, in denen Reimwörter gebildet oder Anlaute erkannt werden müssen.

VORGEHEN

Sprachstandsverfahren müssen so angelegt sein, dass sie jedem Kind, unabhängig von Wohnort oder dem sozialen Hintergrund, Zugang zu einer bedarfsgerechten Sprachförderung ermöglichen. Die Expertenkommission hat daher unter Berücksichtigung ihrer unterschiedlichen Disziplinen, Psychologie, Pädagogik, Sprachdidaktik, Linguistik und Sprachtherapie, erstmals Qualitätsmerkmale formuliert, anhand derer die eingesetzten Sprachstandsverfahren überprüft werden sollten.

Die Sprachstandsverfahren müssen die wesentlichen sprachlichen Kompetenzbereiche erfassen (Aufgabentyp bzw. Beobachtungskategorie), und das unter Bezugnahme auf eine theoretisch und empirisch abgesicherte Basis. Auch müssen diese sprachlichen Bereiche in einer Art und Weise erhoben und analysiert werden, die den Gütekriterien der Messgenauigkeit, Validität und der Objektivität entsprechen.

Ein über die zu berücksichtigenden sprachlichen Bereiche hinausgehendes Merkmal stellt z. B. die Messgenauigkeit des Verfahrens dar. Auch sind in der (Sprach-)Diagnostik zusätzliche Faktoren zu berücksichtigen, so z. B. die Konzentrationsfähigkeit der Kinder, unterschiedliche Zielgruppen (ein- vs. mehrsprachige Kinder etc.), die Qualifikationen und Kompetenzen des Testers bzw. der Testerin, die jeweilige Familiensprache etc.

Darüber hinaus ist es wesentlich, bezüglich dieser Qualitätsmerkmale eines Sprachstandsverfahrens Mindestwerte zu vereinbaren, ab wann ein Verfahren dieses Merkmal hinreichend erfüllt.

WAS IST EIN QUALITÄTSMERKMAL?

Ein Indikator¹, im Folgenden als Qualitätsmerkmal bezeichnet, ist eine Variable, die empirisch ermittelbar ist und anzeigt, wie stark ein referenziertes zugrunde liegendes Merkmal ausgeprägt ist. Ein Indikator muss also mindestens zwei Ausprägungen aufweisen können. Die Entwicklung von Indikatoren stellt einen komplexen sozialen Prozess dar, an dem eine Reihe von Personen mit unterschiedlichem Fachwissen und abweichenden Einstellungen gegenüber dem Messverfahren zu beteiligen sind.

Alle im Folgenden entwickelten Indikatoren werden im ersten Schritt aus dem sprachlichen Kompetenzbereich bzw. aus der Betrachtung notwendiger weiterer Merkmale abgeleitet und dann in eine Form konkreter Tatsachen, die vorliegen (oder auch nicht), operationalisiert.

AB WANN ERFÜLLT EIN SPRACHSTANDS- VERFAHREN EIN QUALITÄTSMERKMAL?

Die Indikatoren liefern ein ausgewähltes, aber objektives Bild der Merkmale von Sprachstandsverfahren. Die Grundlage einer Bewertung ist damit gegeben, für die Bewertung selbst werden aber noch Vereinbarungen benötigt, welche Ausprägung („Indikatorwert“) als „ausreichend“ oder „nicht ausreichend“ bewertet wird². Diese Festlegung von Bewertungskategorien erfolgt für jeden Indikator spezifisch. Als grundsätzliche Abstufungen der Bewertung werden fünf Kategorien durchgängig verwendet, wobei zwei Kategorien in einer Bewertungskategorie zusammengefasst werden: Ist der Indikator für ein Merkmal nicht ermittelbar („unbekannt“), so wird dies automatisch als „nicht ausreichend“ bewertet, da der Nachweis der zentralen Gütemerkmale eines Verfahrens dem Entwickler bzw. dem Ministerium, welches dieses Verfahren vorgibt, obliegt (vgl. Evers, 2001). Die vier Bewertungskategorien sind somit „nicht ausreichend“, „ausreichend“, „gut“ sowie „sehr gut“.

Die klassischen Gütemerkmale Reliabilität, Validität und Objektivität sind hinlänglich im Bereich der Testentwicklung bekannt, allerdings werden diese Anforderungen in Sprachstandsverfahren im Elementarbereich nicht immer berücksichtigt. Dies wird am häufigsten mit einem der zwei folgenden Argumente legitimiert: Das eine Argument ist der Verweis darauf, dass das Verfahren z. B. ein Beobachtungsbogen sei und die genannten Gütekriterien nicht zuträfen. Dies ist insofern falsch, als auch in diesem Fall die zugrunde liegenden Beobachtungen reliabel, valide und objektiv sein müssen, und dies ist – auch bei Beobachtungsbögen – ermittelbar. Zudem wird von sprachwissenschaftlicher Seite angemerkt, dass Gütemerkmale nicht unbedingt aufgrund von Normierungsprozessen erfüllt werden können – vielmehr wird hier eine kriterienbezogene Orientierung an Spracherwerbsprozessen verlangt (Reich, 2007). Als Beispiel für ein solches Vorgehen sind Profilanalysen zu

nennen (vgl. Griebhaber 2005; Heilmann, 2012). Auf der Grundlage der nationalen und internationalen Spracherwerbsforschung werden Erwerbsschritte definiert, die sich in Zweitspracherwerbsprozessen beobachten lassen (z. B. in Bezug auf Satzmodelle). Diese Schritte (auch Meilensteine, vgl. Tracy 2007) sind teilweise relativ altersunabhängig, das heißt ein Kind erwirbt bestimmte Sequenzen mit zwei Jahren, die sich ein anderes Kind erst mit drei Jahren aneignet. Der jeweils nächste Schritt kann zu sehr unterschiedlichen zeitlichen Abschnitten erfolgen. Reine Altersnormen können hier, insbesondere wenn sie den Prozess nicht berücksichtigen, zwar für einen Zeitpunkt valide und reliabel sein, bezogen auf den Zweitspracherwerbsprozess jedoch nicht unbedingt aussagekräftig.

Die Auswahl der Kinder, die Sprachförderung erhalten, ist eine Entscheidung mit weitreichenden bildungsbio-graphischen Konsequenzen. Insofern muss die Entscheidung auf einer soliden Basis erfolgen – das Auswahlverfahren muss den hier dargestellten Standards genügen, während eine Diagnostik des Förderbedarfs die für qualitative Forschung geltenden Standards einhalten muss, insbesondere: intersubjektive Nachvollziehbarkeit und Prüfbarkeit der Ergebnisse; Transparenz; Offenlegung der Instrumentenentwicklung, der Ziele sowie die Veröffentlichung von Erprobungsergebnissen (vgl. Lengyel, 2012, S. 21 ff.)

Das zweite Argument ist der Verweis darauf, dass es ein intern entwickeltes Verfahren sei, dessen Qualität eine „Augenscheinvalidität“³ aufweise und aus der Praxis heraus entwickelt worden sei. Dennoch ist am Anschluss an die (Test-)Entwicklung eine konstruktiv-kritische empirische Analyse vorzunehmen: Welche Stärken und Schwächen weist das so entwickelte Instrument auf? Wie kann das Verfahren darauf aufbauend weiterentwickelt werden? Zumindest die Ergebnisse dieser Analyse sind zu dokumentieren (s. „Informationsverfügbarkeit“).

1 Im „Wörterbuch der Soziologie“ wird unter einem Indikator „eine empirisch direkt (z.B. durch Beobachtung oder Befragung) ermittelbare Größe, die Auskunft gibt über etwas, das selbst nicht direkt ermittelbar ist“ verstanden (Hartmann, 2002). Mayer (2004) definiert Indikatoren wie folgt: „Indikatoren sind Vergleichsinstrumente, die Unterschiede oder Gemeinsamkeiten verschiedener Objekte auf einem gemeinsamen Vergleichsmaßstab (Skala) mit mindestens zwei verschiedenen Ausprägungen (Skalenwerte) anzeigen.“

2 Dies gilt nur für quantitative Daten – qualitative Merkmale können auch dichotom als „erfüllt“ – „nicht erfüllt“ behandelt werden

3 Fachlich betrachtet geht es um die Kontentvalidität. Diese kann systematisch geprüft werden (siehe Klauer 1987)

WAS QUALITÄTSMERKMALE LEISTEN KÖNNEN

Zweck der vorgestellten Qualitätsmerkmale ist zunächst, die im Einsatz befindlichen und für den Einsatz vorgeschlagenen Testverfahren auf ihre Geeignetheit hin zu untersuchen und zu bewerten. Sie sollen die Diskrepanz zwischen Anspruch und Wirklichkeit aufzeigen und den verantwortlichen Akteuren in den Bundesländern konkrete Hinweise für Veränderungen und Anpassungen geben. Bei der Anwendung der Qualitätsmerkmale auf die Sprachstandsverfahren ist zu berücksichtigen, dass vermutlich nicht alle Kriterien vollumfänglich von einem Verfahren erfüllt werden können. Die (Weiter-) Entwicklung von Sprachstandsverfahren im Elementarbereich muss sich aber an klar und eindeutig formulierten wissenschaftlichen Anforderungen orientieren, wie auch die Betroffenen und Beteiligten erwarten können, dass diagnostische Entscheidungen auf der Grundlage rationaler und nachprüfbarer Fakten vorgenommen werden. Hierzu geben die Qualitätsmerkmale möglichst eindeutige Vorgaben, die von jedem Verfahren, insbesondere solchen, die in den Ländern eingesetzt werden, anzustreben und auch weitgehend zu erreichen sind.

EIN BEWERTUNGSRAHMEN FÜR ALLE SPRACHSTANDSVERFAHREN

Die derzeit verwendeten Sprachstandsverfahren werden den drei Begriffen **Beobachtungsbögen**, **Screenings** und **Tests zugeordnet**¹ (vgl. Bildungsbericht 2012; Kany & Schöler 2010). Die drei Arten sind dabei weniger unterschiedlich, als es auf den ersten Blick wirken mag. Erstens sind alle Verfahren dem gleichen Zweck verpflichtet – entweder den Sprachstand eines Kindes möglichst genau zu erfassen, oder aber zumindest, die Kinder mit Förderbedarf auf Grundlage des jeweiligen Sprachstandes möglichst valide, sensitiv und spezifisch zu identifizieren. Bei einem Vergleich von Screenings mit Tests ist eine Besonderheit zu berücksichtigen: Ein Screening weist i. d. R. eine niedrigere Validität und Reliabilität zugunsten der kürzeren Durchführungszeit auf. Durch die ggf. geringere Reliabilität werden systematische und starke Verzerrungen in der Diagnostik verursacht – die teilweise nicht mehr bemerkt werden können.

Bei einem Vergleich von Beobachtungsbögen mit Testverfahren ist ferner zu beachten, dass Tests den gegenwärtigen, punktuellen Zustand untersuchen, während Beobachtungsbögen einen Zeitverlauf berücksichtigen können. Die Datenbasis ist somit bei Beobachtungsbögen „breiter“ und näher am tatsächlichen sprachlichen Verhalten. Beobachtungsbögen – und das ist eine ihrer Stärken – können über die Momentaufnahme hinausgehen und individuelle Lernfortschritte über eine längere Zeitspanne hinweg abbilden. Entsprechend häufig wer-

den sie auch – neben informellen Verfahren – genau für diesen Zweck in der Praxis genutzt.

Allerdings steht diesem Vorteil ein gravierender Nachteil gegenüber: Objektivität ist bei Beobachtungsbögen eine größere Herausforderung als bei einem Test. Wenn eine Beobachtung objektiv sein soll, so ist idealerweise im Vorfeld festzulegen, was beobachtet werden soll. Dies ist in der Praxis nicht der Fall – die Erzieherin bzw. der Erzieher wird sich beim Ausfüllen des Beobachtungsbogens das Kind vorstellen und zu jeder Frage bisherige Erfahrungen mit dem Kind rekonstruieren.²

Trotz dieser Unterschiede können alle hier betrachteten Sprachstandsverfahren, da sie den Zweck der Auswahl für die Sprachförderung verfolgen, mit den gleichen Maßstäben und anhand der gleichen Merkmale erfasst, analysiert und bewertet werden. Eine bessere Verlaufsdagnostik, wie sie v. a. für die Beobachtungsbögen attestiert wird, ist für die hier zu treffende Auswahlentscheidung irrelevant. Ein Screening, welches nur eine dichotome Entscheidung legitimieren muss, benötigt ebenso wie ein Testverfahren die gleiche Basis an Facetten des zu messenden Konstrukts, um eine zuverlässige Entscheidung zu treffen.

Damit werden Beobachtungsverfahren, Screenings und Testverfahren einer gemeinsamen Betrachtung zugänglich: Identifizieren sie die Kinder, die noch weitere Unterstützung bei der Aneignung der (Zweit-) Sprache Deutsch benötigen?

AUSWAHLDIAGNOSTIK VS. DIAGNOSTIK DES FÖRDERBEDARFS

Die Qualitätsmerkmale beziehen sich v. a. auf die Auswahlentscheidung, ob eine Förderung notwendig ist oder nicht, während der Aspekt der differenzierten Diagnostik des Förderbedarfs, der für jede anschließende Sprachfördermaßnahme unabdingbar notwendig ist, als zweiter Schritt betrachtet werden muss. Bei einer Förderung müssen die Sprachförderkräfte so genau wie möglich wissen, wo sie ansetzen. Hierzu wird also eine weitaus differenziertere Diagnose als für die Entscheidung „Förderbedarf: ja oder nein?“ benötigt. Um zu der differenzierteren Diagnostik des Förderbedarfs zu gelangen, muss sichergestellt sein, dass (bei einem zweistufigen Verfahren) die richtigen Kinder in der ersten Stufe ausgewählt werden. Das Verfahren für diese Stufe 1 sollte also, wenn es selbst keine Diagnostik des Förderbedarfs umfasst, klare Konsequenzen beschreiben, worin das weitere Vorgehen besteht.

Ob einzelne oder alle Merkmale der Diagnostik des Förderbedarfs bereits im ersten Schritt integriert werden, ist auch abhängig von übergeordneten Überlegungen: Wenn in einem ersten Schritt zunächst die Kinder ausgewählt werden, die einen besonderen Sprachförderbedarf haben, ist die differenzierte Diagnostik zeitlich weniger aufwändig und so besser in den pädagogischen Alltag integrierbar.

Der Spagat zwischen der Auswahlfunktion und der pädagogischen Förderorientierung ist nur bedingt lösbar – sowohl inhaltlich als auch methodisch. Deshalb wird eine klare Unterscheidung in die zwei Schritte vorgenommen und bei jedem Qualitätsmerkmal explizit angegeben, ob es für die Auswahl oder die Diagnostik des Förderbedarfs relevant ist (vgl. Abbildung 1, Seite 18).

INDIVIDUELLE ENTWICKLUNG VS. NORMALITÄTSERWARTUNGEN

Sprachstandsverfahren müssen die Komplexität des Sprachstandes eines vier- bis fünfjährigen Kindes angemessen erfassen können.³ Es ist plausibel anzunehmen, dass im Elementarbereich einige sprachliche Bereiche relevanter sind als andere – entweder, weil sie besonders gut voraussagen, wie sich die sprachlichen Kompetenzen eines Kindes in der Schule entwickeln werden (prädiktiver Wert)⁴, weil sie nicht beeinflussbar sind oder sich erst mit Beginn der Primarstufe entwickeln (z. B. Schreibkompetenz).

Eine uneingeschränkt akzeptierte Spracherwerbstheorie existiert derzeit nicht (Fried et al. 2007). Dennoch gibt es zahlreiche Fakten, über die ein ausreichender Konsens besteht und für die relativ genau bekannt ist, in welcher Zeitspanne diese Fertigkeit oder Kompetenz von den meisten einsprachigen Kindern üblicherweise erworben wird.⁵

Ein Verfahren kann Kriterien festsetzen, über welche Kompetenzen ein Kind in einem bestimmten Alter verfügen sollte (kriterial). Eine zweite Möglichkeit ist, die sprachliche Entwicklung im Vergleich zu den anderen Kindern einer Gruppe zu untersuchen (normorientiert). Beide Ansätze haben jeweils Vor- und Nachteile, wobei tendenziell kriteriale Vorgaben eher bei Beobachtungsbögen sinnvoll erscheinen, für Screenings und Testverfahren dagegen eine zusätzliche Normierung notwendig ist.

Bei der Betrachtung des Spracherwerbs wird zumeist an einer Norm orientiert verglichen und interpretiert. Normalitätserwartungen sind trotz ihrer hohen gesellschaftlichen Akzeptanz und ihrer faktischen Verbindlichkeit

nicht ohne Schwierigkeiten. Ein zentrales Problem ist dabei die Fixierung bestimmter Etappen der Sprachaneignung auf enge Zeitabschnitte. Ähnliches gilt für den Erwerb anderer Fähigkeiten wie Krabbeln, Stehen, Laufen. Bei der Abweichung von solchen Normalitätsvorstellungen entsteht dann schnell der Ruf nach Interventionen. Vieles davon ist fehl am Platz, weil die individuellen Aneignungen unterschiedliche Geschwindigkeiten haben. Gleichzeitig ist es nicht sinnvoll, von diesen Normalitätserwartungen komplett abzuweichen: Es ist sinnvoll, sich an dem „ungestörten, sich erfolgreich entwickelnden“ Spracherwerb zu orientieren, um bei Abweichungen entscheiden zu können, ob diese sich im Laufe der weiteren Entwicklung ohne Intervention nivellieren oder ob eine gezielte Intervention notwendig ist, da ansonsten ein Zeitfenster der Entwicklung ggf. verpasst wird. Auch orientiert sich die Schule an diesen Normalitäts-

erwartungen – ein Kind, das bei der Einschulung diesen Erwartungen nicht ausreichend genügt, wird dem Unterricht nicht angemessen folgen können. Dass die Lehrkräfte diese Heterogenität in ihrem Unterricht mit 20 bis 30 Kindern didaktisch problemlos auflösen können, erscheint wünschenswert, aber wenig realistisch.⁶

Das bedeutet, die Normalitätsvorstellungen der Schule als vorerst gegeben anzusehen. Die Verantwortung der Institution Schule, sich weiterzuentwickeln – auch in diesem Bereich – darf aber nicht ignoriert werden. Es hat sich gezeigt, dass schulische Normalitätsvorstellungen häufig unrealistisch sind, wenn vom typischen Fall eines „normalen“ Erwerbs einer Verkehrssprache ausgegangen wird, was der Realität einer multikulturell geprägten Gesellschaft nicht gerecht wird.

MEHRSTUFIGES DIAGNOSEVERFAHREN

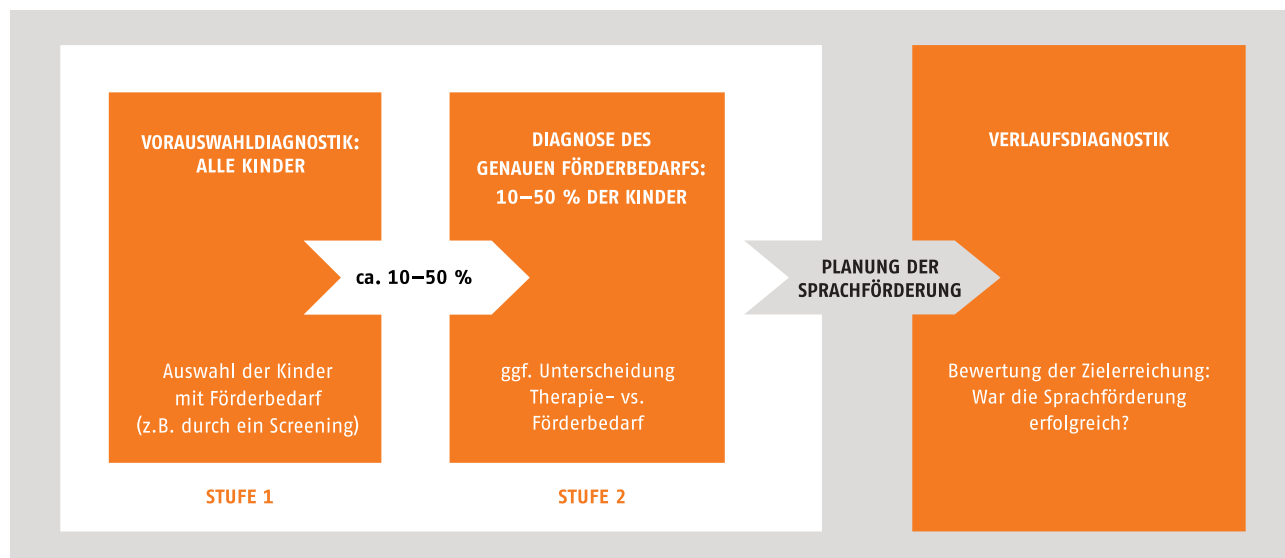


Abbildung 1: Mehrstufigkeit des diagnostischen Prozesses

1 Spontansprachanalysen oder die Analyse freier Sprachproben, häufig als „Profilanalyse“ bezeichnet (z. B. Heilmann, 2012), werden bei diesen Kategorisierungen den „Beobachtungsbögen“ zugerechnet, so z.B. HAVAS 5.

2 Dieses Vorgehen entspricht einem Ratingverfahren zur Fremdbeurteilung.

3 Diese Voraussetzung wird als Konstruktvalidität bezeichnet und zu den klassischen Gütemerkmalen gerechnet.

4 Z. B. für den Schriftspracherwerb die phonologische Bewusstheit, vgl. Hasselhorn, Schneider & Marx (2000), siehe aber auch Valtin (2012),

für die Lesesozialisation die Literalität (z. B. Ulich, 2003b), im Englischen als Literacy bezeichnet.

5 Derzeit liegen überwiegend qualitative, nicht aber quantitative Untersuchungen zu dem Erwerbsalter klar umrissener Teilfertigkeiten oder –kompetenzen vor. Eine detaillierte Übersicht findet sich z. B. im Referenzrahmen zur altersspezifischen Sprachaneignung (Ehlich, Bredel & Reich, 2008).

6 „Die Orientierung an der Altersnorm ist allerdings die Realität an den Schulen, in denen sich mehrsprachige Kinder mit einsprachigen vergleichen lassen müssen, ohne dass auf ihre spezifischen Bedürfnisse eingegangen wird.“ (Jeuk 2009, Abschnitt 3)

QUALITÄTSMERKMALE

BERÜCKSICHTIGUNG SPRACHLICHER BASISQUALIFIKATIONEN

Das Verfahren erfasst die zentralen Bereiche der sprachlichen Entwicklung eines Kindes.

Im vorliegenden Ansatz werden sieben der acht Sprachbereiche des Qualifikationsfächers nach Ehlich (2007, S. 14) für den Elementarbereich als relevant betrachtet.^{1,2} Die Basisqualifikationen können nach Reich (2008, S. 165) „[...] auf alle menschlichen Lautsprachen angewendet werden, finden also auch bei der Mehrsprachigkeitsdiagnostik Verwendung“.

Die sieben ausgewählten Basisqualifikationen sind einerseits empirisch gut ermittelbar, gleichzeitig weisen sie nach derzeitigem Kenntnisstand eine hohe prognostische Validität für den erfolgreichen Erwerb der schulbezogenen Bildungssprache auf.³ Allerdings sind die ausgewählten Basisqualifikationen mit verschiedenen Verfahrenstypen unterschiedlich gut bestimmbar – so kann die diskursive Basisqualifikation leicht mit einem Beobachtungsverfahren eingeschätzt werden, dagegen sind – zumindest ausgewählte – Aspekte der semantischen Basisqualifikation problemlos mit einem Screeningverfahren bestimmbar.

Phonische Basisqualifikation

Sie umfasst die Wahrnehmung, Erfassung und Produktion von Lauten, Silben und größeren Einheiten sowie von übergreifenden zielsprachlichen intonatorischen Strukturen. Der Aussprache von Kindern liegt ein auf phonologischen Prozessen basierendes Regelsystem zugrunde. Durch die stufenweise Überwindung dieser Prozesse erweitern sie ihr Aussprachesystem und nähern sich den Regeln der Erwachsenensprache an.

Der Erwerb der Lautbildung vollzieht sich nach derzeitigem Kenntnisstand stufenweise. Eine Sprachförderung ist notwendig, wenn zeitliche Verzögerungen bei der Überwindung von phonologischen Prozessen oder beim Erwerb des Lautsystems auftreten. Eine Sprachtherapie ist dann erforderlich, wenn ergänzend zu den zeitlichen Verzögerungen ungewöhnliche phonologische Prozesse auftreten, z. B. eine „Lautpräferenz“ oder eine „unausbalancierte Entwicklung“ (Hacker 2002, S.32). Dies führt in der Regel dazu, dass das Lautinventar eingeschränkt ist.

Die pragmatische Basisqualifikation I

Sie bildet die kommunikative Grundlage für die Aneignung von Sprache. Sie wird zunächst in der frühen Interaktion des Kindes mit seinen engsten Bezugspersonen ausgebildet. Das Kind lernt, aus dem Einsatz von Sprache der anderen Interaktionspartner deren Handlungsziele zu erkennen; es eignet sich zugleich an, wie es selbst Sprache zum Erreichen eigener Handlungsziele angemessen einsetzen kann. Mit zunehmendem Alter lernt das Kind, zwischen seiner eigenen Perspektive und der des Interaktionspartners zu unterscheiden und sein sprachliches Handeln dementsprechend zu gestalten.

Semantische Basisqualifikation

Sie betrifft die Wörteraneignung, darüber hinaus aber auch die Begriffsbildung und die Übertragung von Bedeutungen sowie die Ermittlung von Satzbedeutungen, die durch das Zusammenspiel von Wortbedeutungen und der Kombination von Wörtern im Satz entstehen – die Zuordnung sprachlicher Ausdrücke zu Wirklichkeitselementen und zu Vorstellungselementen. Als einfachster Indikator dieser Qualifikation kann eine Erfassung des Wortschatzes bzw. des Wortschatzerwerbs dienen, darüber hinaus sollten aber auch die Begriffsbildung und die Sprachbetrachtung, das Spielen mit Wörtern und ihren Bedeutungsmöglichkeiten, berücksichtigt werden. Zentral ist, dass Kinder die Fähigkeit erwerben, ihre Begriffsbildung und ihren Wortschatz selbstständig zu erweitern, indem sie z. B. nach Bedeutungen fragen.

Morphologisch-syntaktische Basisqualifikation

Für die deutsche Sprache gilt, dass sich das Kind im Verlauf der ersten sechs Lebensjahre wesentliche Kompetenzen aneignet, und zwar im Bereich der Syntax (z. B. Stellung von Wörtern in einer Äußerung) sowie im Bereich der Morphologie (Bildung neuer Wörter (Kind – kindlich) und Bildung von Wortformen durch Flexion (Beugung) (komm/st)).

Die diskursive Basisqualifikation

Sie betrifft zunächst die grundlegenden Strukturen der formalen sprachlichen Kooperation (Verfahren des Sprecherwechsels), ohne die Kommunikation als (zumindest) dyadische Interaktion nicht gelingen kann. Diese Verfahren werden bereits im ersten Lebensjahr vorsprachlich eingeübt und bis ins Schulalter hinein ausdifferenziert (Organisation des Sprecherwechsels in Institutionen). Die diskursive Basisqualifikation erfasst neben der Aneignung dieser eher formalen Strukturen die Befähigung zum komplexen zweckgerichteten sprachlichen Handeln mit anderen. Einen zentralen

Teilbereich der diskursiven Basisqualifikation stellt die Aneignung zuhörerorientierter Erzählfähigkeiten dar, die mit etwa drei Jahren beginnt und bis ins Schulalter hinein weiterentwickelt wird.

Die pragmatische Basisqualifikation II

Sie erfasst diejenigen pragmatischen Kompetenzen des Kindes, die mit Eintritt in eine Bildungsinstitution (i. d. R. zunächst die Kinderkrippe oder Kindertagesstätte, später die Schule) relevant werden. Die Aneignung angemessener sprachlicher Mittel für den Einsatz in unterschiedlichen sozialen Wirklichkeitsbereichen stellt für das Kind einen wichtigen Schritt in seiner sprachlichen Entwicklung dar. Umgekehrt muss mit erheblichen Schwierigkeiten gerechnet werden, wenn eine solche Aneignung nicht gelingt.

Die literale Basisqualifikation I

Hier ist auch die bereits erwähnte Literalität anzusiedeln. Wenn im Elternhaus wenig (vor)gelesen wird, unabhängig davon, in welcher Familiensprache, fehlt diesen Kindern ein wesentlicher Baustein der Literacy-Entwicklung. Sie haben keine Textsorten kennengelernt und verfügen nur über eingeschränkte Erfahrungen mit Geschichten im weitesten Sinne. Die diskursive Basisqualifikation ist mit der Herausbildung der literalen Basisqualifikation I (und später dann auch II) naturgemäß verbunden, sofern das literarische Erzählen auf dem alltäglichen Erzählen aufbaut und es weiterentwickelt. Der Übergang in die Welt der Schrift ist aber eine eigene, nicht zuletzt auch für die Entwicklung des Kindes sehr wichtige Epoche seiner eigenen sprachlichen Qualifizierung.

Sie ist insofern wichtig, als im Alltagsverständnis die Beschäftigung mit dem Erwerb der Schrift beim Eintritt in die Schule beginnt. Doch die Beschäftigung mit Schrift ist ein Teil der sprachlichen Bildung von Anfang an – allerdings nur in Familien, in denen sie auch Bedeutung hat. Es ist nachgewiesen, dass Kinder aus bildungsfernen Elternhäusern und viele Kinder mit Migrationshintergrund ohne Bücher und fast ohne literarische Erfahrungen aufwachsen. Vor allem diese Kinder benötigen eine frühe Unterstützung.⁴

Somit sind die explizite Ausweisung und Anzahl der berücksichtigten Basisqualifikationen wie auch die Differenziertheit der Aufgaben, d. h. eine Basisqualifikation mit unterschiedlichen Aufgabentypen zu erfassen, relevant. So nennt Lisker (2011): „Die Sprachebene der Phonetik-Phonologie wird in zehn der 16 Verfahren erfasst, wobei sich die meisten Verfahren darauf beschränken,

die Artikulation einzuschätzen“ (S. 55). Eine genauere Einschätzung ist dann gegeben, wenn die Phonologie über mehrere unterschiedliche Beobachtungsbereiche oder Aufgabentypen abgedeckt wird.

Auch muss für diesen Altersbereich eine Berücksichtigung instabiler Übergangserscheinungen erfolgen (vgl. Lengyel 2012, S. 21), dass also z. B. grammatische Fertigkeiten für eine kurze Zeit korrekt verwendet werden,

darauf folgend aber z. B. durch eine Übergeneralisierung des neu erworbenen Grammatikwissens mehr Fehler gemacht werden als zuvor. Als letzter Punkt muss die Validität der Aufgabe, also der Bezug zum theoretischen Konstrukt der Basisqualifikation, sichergestellt sein: „Dies beruht größtenteils auf Plausibilitätsabwägungen und wird nur selten hinreichend belegt.“ (Lisker, 2011, S. 53).

QUALITÄTSMERKMALE SPRACHLICHE BASISQUALIFIKATIONEN

In Bezug auf die sprachlichen Basisqualifikationen müssen Sprachstandsverfahren im Elementarbereich folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Ausweisung und Anzahl der berücksichtigten Basisqualifikationen	Keine Angaben; weniger als 3 Basisqualifikationen werden berücksichtigt.	Basisqualifikationen werden benannt, mind. 3 werden berücksichtigt	Basisqualifikationen werden benannt, mind. 4 werden berücksichtigt.	Basisqualifikationen werden benannt, mind. 6 werden berücksichtigt.
Differenziertheit der Aufgaben	Basisqualifikation wird mit einer einzelnen Aufgabe / einem Item abgefragt.	Eine Basisqualifikation wird mit mind. 2 Aufgaben erfasst.	Eine Basisqualifikation wird mit mind. 2 unterschiedlichen Aufgabentypen erfasst.	Eine Basisqualifikation wird mit mind. 3 unterschiedlichen Aufgabentypen erfasst.
Berücksichtigung instabiler Übergangserscheinungen	Keine Angaben	Bei der Interpretation der Ergebnisse wird auf ggf. bestehende instabile Übergangserscheinungen hingewiesen.	Konkrete sprachliche Bereiche, die instabile Übergangserscheinungen aufweisen, werden benannt.	Der Einfluss instabiler Übergangserscheinungen auf die Ergebnisse wird in den einzelnen Bereichen ausgewiesen.
Bezug zur Basisqualifikation: über ein Außenkriterium numerisch-quantifiziert	Keine Angaben; Validität unter 0.6	Validität des Aufgabentypus für die Basisqualifikation ist $r \geq 0.6$.	Validität des Aufgabentypus für die Basisqualifikation ist $r \geq 0.7$.	Validität des Aufgabentypus für die Basisqualifikation ist $r \geq 0.8$.
Bezug zur Basisqualifikation: über ein Expertenrating mit einer Interrater-übereinstimmung (als Cohens Kappa) ⁵	Keine Angaben; wird nur von TestentwicklerIn attestiert.	Von mind. 3 Fachpersonen bewertet; Cohens Kappa ≥ 0.4	Von mind. 5 Fachpersonen bewertet; Cohens Kappa ≥ 0.6	Von mind. 7 Fachpersonen bewertet; Cohens Kappa ≥ 0.8

Tabelle 3

VALIDITÄT

Das Verfahren ist so konstruiert, dass es exakt den Sprachstand eines Kindes erfasst und nicht andere Bereiche, wie zum Beispiel kulturelles Wissen.

Um die sprachliche Entwicklung eines Kindes zu untersuchen, müssen die einzelnen sprachlichen Kompetenzen mit den richtigen Fragen und Aufgaben überprüft werden. Je besser dies gelingt, umso valider ist das Instrument. Die Erfassung des kindlichen Sprachstandes kann über sehr unterschiedliche Umsetzungen erfolgen – über die Beobachtung, Einschätzung, oder auch durch das Lösen vorgegebener Aufgaben. Während die beiden

erstgenannten Verfahren keine Reaktanz auslösen können, ist bei einer responsiven Erfassung/Beobachtung darauf zu achten, das Kind nicht in eine atypische Situation zu bringen und damit nicht das kindliche Sprachniveau zu erfassen, sondern die Fähigkeit des Kindes, mit der ungewohnten Situation umzugehen.

Die externe Validität, auch Allgemeingültigkeit, Verallgemeinerungsfähigkeit oder ökologische Validität, bezeichnet die Übereinstimmung von tatsächlichem und intendiertem Untersuchungsgegenstand. Es ist anzustreben, dass die Diagnose in einer alltagsintegrierten Situation stattfindet, d. h. in Alltags- und Spielformaten und nicht in einer atypischen Testsituation.

Wird die Validität – üblicherweise in Form des Korrelationskoeffizienten berichtet – betrachtet, ergeben sich folgende Anforderungen:

QUALITÄTSMERKMALE VALIDITÄT

Ein valides Sprachstandsverfahren muss folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Bezug zu eindeutig benannter Spracherwerbstheorie	Keine Angaben; „aktuelle Spracherwerbstheorien“	Wird eindeutig mit Quelle benannt	Aufgaben sind aus spezifischen Aussagen der Theorie abgeleitet.	Güte der Ableitung (Operationalisierung) wird angegeben.
Konstruktvalidität ist von Fachpersonen nachvollziehbar.	Keine Angaben; unbekannt; Wird von Testautoren attestiert.	3 externe Fachpersonen aus unterschiedlichen Fachbereichen (z. B. Sprachwissenschaften, Fachdidaktik, Pädagogik, Psychologie) bescheinigen übereinstimmend „hohe“ Konstruktvalidität.		
Empirische Faktorenstruktur	Keine Angaben; unbekannt	Faktorenstruktur wird mitgeteilt.	Faktorenstruktur explorativ bestätigt	Faktorenstruktur konfirmatorisch gültig
Prognostische Validität	Keine Angaben; unbekannt	RATZ-Index > 30	RATZ-Index > 40	RATZ-Index > 65
Konvergente oder kriterienorientierte Validität ⁶	Keine Angaben; unbekannt	$r^2 > 50 \%$	$r^2 > 60 \%$	$r^2 > 70 \%$; diskriminante Validität wird mitgeteilt.
Alltagsnähe der Testsituation	Testsituation / Beobachtung verwirrt das Kind / verängstigt / setzt es unter Stress.	Die Testsituation setzt das Kind nicht unter Stress.	Das Kind kann sich in der Beobachtungs- oder Testsituation frei verhalten.	Das Kind erlebt die Beobachtungs- oder Testsituation als positiv.
Alltagsnähe der Aufgaben / Beobachtungssituation	Die Aufgaben / Beobachtungssituation entspricht in keiner Weise den Alltagserfahrungen des Kindes.	Die Aufgaben / Beobachtungen werden in einer dem Kind verständlichen Weise umgesetzt.	Das Kind kann die Test- / Beobachtungssituation verstehen.	Das Kind findet die Test- / Beobachtungssituation anregend / bewusstseinsbildend.

Tabelle 4

Hinweis: RATZ = Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote (Marx 1992)

OBJEKTIVITÄT

Das Verfahren ist unabhängig von subjektiven Einflussfaktoren, wie beispielsweise dem Verhalten der pädagogischen Fachkraft.

Die Standardisierung eines Verfahrens ist dann optimal, wenn das Verhalten des Kindes in der Testsituation die einzige Variationsquelle darstellt, alle anderen Bedingungen hingegen konstant oder kontrolliert sind, so dass sie nicht als Störvariablen wirken können. Die Leistung soll also nur von der Merkmalsausprägung des Individuums abhängen.⁷ Es sind die Durchführungs-, Auswertungs- und Interpretationsobjektivität zu unterscheiden (z. B. Lienert & Raatz, 1998).

Durchführung

Es treten nur systematisch geplante sprachliche Interaktionen auf, die vielfältigen kindlichen Reaktionen sind durch ein Manual abgedeckt. Die Durchführung verläuft weitgehend störungsfrei. Mögliche Verzerrungen (bewusst oder unbewusst herbeigeführt) durch Testleiterinnen und -leiter sind weitestgehend ausgeschlossen.

Auswertung

Die beobachteten sprachlichen Reaktionen können eindeutig einer Beobachtungskategorie/einer Aufgabenstellung zugeordnet werden. Die Auswertungskategorien sind eindeutig definiert, Kategorien wie „häufig“ oder „selten“ sind eindeutig beschrieben.

Interpretation

Das beobachtete Verhalten kann in Bezug gesetzt werden zu einem kriterien- oder sozialnormierten Referenzrahmen. Hier sind auch Angaben enthalten, welche Ausprägung über- und unterdurchschnittlich zu bewerten und ab welchem Wert ein Defizit oder eine Störung zu diagnostizieren ist. Die Interpretationsanweisung macht die gesellschaftlichen/normativen Erwartungen explizit. Bei der Anleitung für die Auswertung und Interpretation ist anzugeben, wie hoch die Messgenau-

igkeit ist, um eine Fehlinterpretation oder eine Überbewertung von Unterschieden auszuschließen.

Die Objektivität ist eine notwendige Voraussetzung für die Messgenauigkeit und die Validität, da beide Kennwerte nur bei Einhaltung der Objektivität hoch sein können. Wenn eine geringe Objektivität vorliegt, sagt das Ergebnis eines Verfahrens mehr über die Testleitung als über das Kind aus. Sowohl Wirtz und Caspar (2002) als auch Evers (2001) schlagen hierbei Untergrenzen der Beobachter- bzw. Beurteilerübereinstimmung vor. Bei dichotomen Bewertungen (z. B. Förderbedarf: Ja vs. Nein) empfehlen Wirtz und Caspar (2002) die Berechnung der Objektivität durch die zufallskorrigierte prozentuale Übereinstimmung (RATZ-Index, z. B. Marx & Weber, 2006, Marx & Lenhard, 2010) oder Cohens Kappa. Im Folgenden wird Cohens Kappa verwendet, um Untergrenzen einer ausreichenden Objektivität zu bestimmen. Werden bei einem Sprachstandsverfahren andere statistische Kennwerte zur Ermittlung der Objektivität angegeben, so sind diese in der Regel problemlos transformierbar.

Hinsichtlich des Verfahrens selbst wird durch das Manual (als essenzieller Bestandteil des Verfahrens) abgesichert, dass der Ablauf des diagnostischen Prozesses durch ausreichend qualifizierte Personen kompetent, nachvollziehbar und eindeutig durchgeführt werden kann. Die sprachlichen Handlungen müssen bei Testverfahren und Screenings als „standardisierte Interaktionen“ umgesetzt werden, um eine Einheitlichkeit der Durchführung zu gewährleisten.

Optimal wird die Objektivität z. B. im Bereich der Durchführung umgesetzt, wenn

... bei einem Test oder Screening der Testleiter exakt den vorgegebenen Stimulus wiedergibt und alle sich daran anschließenden sprachlichen Interaktionen durch das Manual abgedeckt werden;

... bei einem Beobachtungsverfahren der Testleiter bereits im Vorfeld des Ausfüllens des Bogens das zu erfassende Verhalten in systematisch ausgewählten Situationen bei den Kindern beobachtet.

QUALITÄTSMERKMALE OBJEKTIVITÄT

Ein objektives Sprachstandsverfahren muss folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Manual / Testheft des Verfahrens	Kein Manual	Es liegt ein Manual mit eindeutigen Hinweisen zur Durchführung, Auswertung und Interpretation vor.	Die sprachlichen Interaktionen werden vollständig durch das Manual standardisiert.	Das Manual wird flankiert durch Schulungen und externe Beratung.
Objektivität	Die Objektivität ist unbekannt / liegt unter 0.9.	Die Objektivität liegt zumindest in einem Bereich bei mind. 0.9 und die Durchführungsobjektivität liegt bei mind. 0.8.	Die Objektivität liegt global / als Gesamtangabe bei mind. 0.9.	Die Objektivität liegt in allen Bereichen über 0.9.

Tabelle 5

RELIABILITÄT

Das Verfahren misst den Sprachstand des Kindes präzise. Wird das Verfahren wiederholt, steht am Ende das gleiche Ergebnis.

Durchgängig wird die Messgenauigkeit bei diagnostischen Verfahren eingefordert. Obwohl als Reliabilität in quantifizierbarer Form vorliegend, wurden im Bereich der Sprachdiagnostik für den Elementarbereich bislang keine spezifischen Anforderungen formuliert. Da es sich um Individualdiagnostik mit Bedeutung für die individuelle Biographie handelt, ist eine höhere Reliabilität erforderlich, als wenn es sich um die Untersuchung z.B. eines Förderprogrammes handelt, da im letzteren Falle Gruppenmittelwerte gebildet werden und eine individuelle Falschklassifikation ohne Konsequenzen für das Individuum bleibt.

In Anlehnung an das COTAN-System (Evers, 2001) kann eine Untergrenze für die Messgenauigkeit definiert wer-

den. Während Evers (2001) allerdings den allgemeinen Fall der Testdiagnostik betrachtet, sind hier die höhere individuelle kindliche Schwankungsbreite und Erwerbsdynamik zu berücksichtigen, wie sie u.a. aus dem teilweise nichtlinearen Erwerb von neuen sprachlichen Kompetenzen / Fähigkeiten resultiert. Dies berücksichtigend wird das Kriterium für „ausreichend“ um den Wert 0.1 verschoben.

Hinsichtlich der Reliabilität wird gefordert, dass diese als Retest- oder als Parallelfarm-Reliabilität erhoben und berechnet wird. Andere Formen der Reliabilitätsberechnung (z.B. Split-Half, Interne Konsistenzen) erscheinen zumindest im Rahmen der Gesamtbetrachtung eines Verfahrens nicht sinnvoll. Da das Konstrukt des Sprachstandes als Zusammenschluss verschiedener Qualifikationen gesehen wird, ist es sehr unwahrscheinlich, dass ein Test, der all diesen oben geforderten Kategorien genügt, tatsächlich eine hohe Reliabilität aufweist. Sinnvoll wäre in diesem Fall, dass man für einzelne Qualifikationen unterschiedliche Skalen in einem Test hat. Hier könnte dann wieder die Reliabilität der einzelnen Skalen gemessen werden.

COTAN-SYSTEM (EVERS 2001):

Beispiel für die Richtlinien zur Bewertung der Reliabilität sowie des Umfanges der Normstichproben

	Reliabilität			Umfang Normstichprobe		
	Niveau 1	Niveau 2	Niveau 3	Niveau 1	Niveau 2	Niveau 3
Unzureichend	< .80	< .70	< .60	< 300	< 200	< 100
Ausreichend	.80 – .90	.70 – .80	.60 – .70	300–400	200–300	100–200
Gut	> .90	> .80	> .70	> 400	> 300	> 200

Tabelle 6 (Anmerkungen: Niveau (1) Tests für wichtige Entscheidungen auf individueller Ebene (z. B. Personalauswahlentscheidungen), (2) Tests für weniger bedeutsame Entscheidungen auf der individuellen Ebene (z. B. Fortschrittskontrolle), (3) Tests für Untersuchungen auf Gruppenniveau. Aus: Kersting 2006, Testkuratorium.)

QUALITÄTSMERKMALE RELIABILITÄT

Ein reliables Sprachstandsverfahren muss folgendes Kriterium erfüllen:

	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Reliabilität des Verfahrens	Unbekannt; $r < 0.7$	$0.7 < r < 0.8$	$0.8 < r < 0.9$ oder differenziert nach Subskala > 0.7	$r > 0.9$ oder differenziert nach Subskala > 0.8

Tabelle 7

NORMIERUNG

Die Ergebnisse eines Kindes lassen sich mit den Ergebnissen anderer Kinder vergleichen. Nur so lässt ein Verfahren eine Aussage darüber zu, welche Kinder einen Sprachförderbedarf haben und welche Kinder sich sprachlich altersgerecht entwickeln.

Mit der Normierung wird der Vergleich von Ergebnissen eines Kindes mit den Ergebnissen anderer Kinder ermöglicht. Rohwerte als Ergebnis von Verfahren werden

üblicherweise in einen Normwert transformiert und erlauben erst dadurch eine Bewertung des Ergebnisses. Die Ermittlung dieser Norm anhand der Normierungsstichprobe wird in der Testkonstruktion Normierung genannt. Die Normierung ist zentral, da diese die oben erläuterten „Normalitätserwartungen“ spezifizieren. Je größer und repräsentativer die Normierungsstichprobe, um so genauer kann der ermittelte Wert verglichen werden. Es sollte deshalb vor der Durchführung einer Untersuchung entschieden werden, wie viele Personen benötigt werden, um Aussagen mit der gewünschten Genauigkeit machen zu können. Damit dieser Vergleich mit einer Normstichprobe zulässig ist, sind folgende Kriterien zu erfüllen:

QUALITÄTSMERKMALE NORMIERUNG

Für eine zulässige Einordnung der Ergebnisse muss ein Sprachstandsverfahren folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Bezug zu den Teilbereichen	Keine Angaben	Es wird darauf hingewiesen, welche Bereiche ausreichend normiert sind.	Die Normierung ist (auch) für die Teilleistungsbereiche spezifisch.	Wie „gut“ sowie Angabe der spezifischen Konfidenzbänder
Alter der Normierung	Unbekannt; Normierung älter als 12 Jahre	Normierung max. 12 Jahre alt	Die Normierung ist nicht älter als 8 Jahre (vgl. DIN 33430).	Normierung wird jährlich / alle 2 Jahre aktualisiert.
Umfang der Normstichprobe	N < 200	N > 200	N > 250	N = 300
Art der Gewinnung der Normstichprobe	Unbekannt	Stichprobengewinnung wird beschrieben, keine Randomisierung.	Stichprobengewinnung wird beschrieben, randomisierte Auswahl, allerdings liegt eine Klumpenstichprobe vor, da Daten in Gruppen (ganze Kindertagesstätten) erhoben wurden.	Stichprobengewinnung wird beschrieben, vollständig randomisierte Auswahl auf Ebene der Individuen.
Unterscheidung unterschiedlicher Zielgruppen	Wird nicht vorgenommen; unbekannt, für wen die Normierung gültig ist.	Die Normierungsstichprobe wird detailliert dargestellt.	Die Normierungsstichprobe wird für alle Zielgruppen des Verfahrens detailliert dargestellt.	

Tabelle 8

FEHLERQUOTE (GEWICHTUNG DER AUSWAHLFEHLER)

Das Verfahren stellt sicher, dass die Kinder, die einen Sprachförderbedarf haben, auch wirklich identifiziert werden.

Kein Test- oder Beobachtungsverfahren kann perfekt sein. Es wird immer vorkommen, dass Kinder, die keinen Förderbedarf haben, als förderbedürftig klassifiziert werden. Auch wird es vorkommen, dass Kinder, die Förderbedarf aufweisen, irrtümlich als „nicht förderbedürftig“ klassifiziert werden. Um diese Falschklassifikationen zu mindestens zu minimieren, wurden Obergrenzen entwickelt.

Da es jedoch – wie genannt – kein Testverfahren es leisten kann, eine fehlerfreie Auswahl zu treffen, ist offen, welche Falschklassifikation negativere Folgen hat: Kinder ohne Förderbedarf zu fördern, oder Kinder mit Förderbedarf nicht zu erkennen. Auch hängt diese Fragestellung von weiteren Faktoren ab, wie z. B. dem Kriterium, ab wann ein Kind einen Förderbedarf aufweist, wie auch vom zur Verfügung stehenden Budget für die Sprachförderung.

In der Diagnostik werden der sog. Alpha- und der Beta-Fehler unterschieden und als Sensitivität und Spezifität eines Verfahrens bezeichnet⁸. Hierbei ist zu beachten, dass die Falschklassifikationen auch von der Höhe der Reliabilität abhängen – je höher die Reliabilität, desto weniger Falschklassifikationen werden vorgenommen. Die folgende Tabelle veranschaulicht diese Unterscheidung.

	„Wahrheit“: Förderbedarf	„Wahrheit“: kein Förderbedarf
Testergebnis: Förderbedarf	Korrekt (Test ist sensitiv)	Alpha- Fehler
Testergebnis: kein Förderbedarf	Beta-Fehler	Korrekt (Test ist spezifisch)

Tabelle 9

Bei „preschool language tests“ wird z. B. von Plante und Vance (1994) eine Sensitivität und Spezifität von mind. 90 Prozent gefordert, ein Wert von unter 80 Prozent gilt als „unzureichend“. Eine fehlende Förderung von Kindern mit Förderbedarf (Beta-Fehler) hat allerdings gravierendere Konsequenzen als die Förderung von Kindern, die eigentlich keinen Förderbedarf haben. Ein zweites Argument für eine höhere Gewichtung des Beta-Fehlers (Sensitivität) ergibt sich aus dem vorgeschlagenen zweistufigen Verfahren – bei einer Fehlklassifikation eines Kindes mit normaler Sprachentwicklung würde die Falschklassifikation in der ersten Stufe mit hoher Wahrscheinlichkeit in Stufe 2 entdeckt werden. Entsprechend sollte die Sensitivität höher liegen als die Spezifität. Als Durchschnittswerte für Sprachscreenings haben z.B. Law, Boyle, Harris, Harkness und Nye (2000) einer Sensitivität von 78 Prozent und einer Spezifität von 88 Prozent ermittelt.

Die Sensitivität und Spezifität sind evtl. nicht geeignet, da die Bewertungsmaßstäbe sehr davon abhängig sind,

wie groß die Auffälligkeitsrate ist (diese liegt, wie genannt, zwischen zehn und 50 Prozent). Alternativ kann der RATZ-Index eingesetzt werden, um die jeweiligen Fehldiagnosen anzugeben. Janssen et al. (1999) geben an, dass Werte über 66 Prozent als sehr gute Klassifikationen bezeichnet werden, Werte zwischen 34 und 66 Prozent gelten als gute (aber eher als nicht spezifische) Klassifikationen und Werte unter 34 Prozent als nicht akzeptabel.

Der RATZ-Index setzt die Trefferquote in Bezug zur Zufallstrefferquote und zur maximal möglichen Trefferquote: Sobald die Grundrate von Kindern mit Förderbedarf und die Anzahl geförderter Kinder nicht übereinstimmen, liegt die maximal mögliche Trefferquote unter 100 Prozent. Diesen Umstand berücksichtigt der RATZ-Index, indem der Anstieg der Trefferquote gegenüber der Zufallstrefferquote in Relation zur maximal möglichen Trefferquote betrachtet wird: Der RATZ-Index gibt an, wie deutlich das Verfahren den Zufall „schlägt“.

QUALITÄTSMERKMALE AUSWAHLFEHLER

Für eine Minimierung der Auswahlfehler müssen Sprachstandsverfahren folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Sensitivität	Unbekannt; < 60 %	60 % – 69 %	70 % – 80 %	> 80 %
Spezifität	Unbekannt; < 50 %	50 % – 59 %	60 % – 70 %	> 70 %
RATZ – Index Identifikation „normaler“ Kinder	Unbekannt; < 35 %	35 % – 49 %	50 % – 65 %	> 65 %
RATZ – Index Identifikation förderbedürftiger Kinder	Unbekannt; < 40 %	40 % – 54 %	55 % – 70 %	> 70 %

Tabelle 10

QUALIFIZIERUNG DER PÄDAGOGISCHEN FACHKRÄFTE

Die Personen, die den Sprachstand der Kinder messen, sind im Vorfeld für Sprachdiagnostik sensibilisiert und qualifiziert worden.

Das Sprachstandsverfahren an sich gewährleistet noch nicht Qualität. Erst wenn es fachlich kompetent durchgeführt wird und die zeitlichen Ressourcen der Beteiligten berücksichtigt werden, ist die Qualität des diagnostischen Prozesses gesichert. „Die Güte des besten Verfahrens ist nie höher als die Kompetenz der Untersucherin zu seiner Durchführung“ (Kany & Schöler 2010, S. 118).

Die Qualifikation pädagogischer Fachkräfte zur Diagnostik ist derzeit nicht ausreichend gegeben (vgl. Schöler et al. 2011). Ähnliche Befunde existieren für Lehrkräfte der Primarstufe. Je nach Komplexität und dem Grad der Standardisierung werden entsprechend unterschiedlich häufig Fehler bei der Durchführung, Auswertung und Interpretation passieren – je weniger standardisiert und je komplexer das Verfahren, desto häufiger werden systematische und unbemerkte Fehler in die Diagnostik einfließen.

Unbemerkte Fehler sind doppelt kritisch, weil die diagnostizierende Person überzeugt sein wird, alles richtig zu machen (woraus eine Abwehrhaltung gegenüber Kritik resultieren kann). Solche systematischen Fehler sind nicht mehr erkennbar, wenn die Ergebnisse an Dritte übermittelt werden. Entsprechend wichtig ist es, dass

die benötigten Qualifikationen vorliegen. Die Testerinnen und Tester müssen nicht nur das technische Manual, sondern auch die Grundlage der Diagnostik verstehen. Nur so kann sichergestellt werden, dass der Test inhaltlich kritisch reflektiert wurde.

Zudem ist insbesondere der Bereich der Auswertung nicht trivial – so geben immerhin 15 Prozent der Förderkräfte an, dass sie bei der Auswertung (sehr) unsicher seien (Neugebauer 2010). Diese Herausforderung ist mehrfach auch für den deutschsprachigen Raum thematisiert worden. Trotz teilweise sehr kontrovers geführter Diskussionen über die Notwendigkeit und Berechtigung der Erweiterung des Auftrages an pädagogische Fachkräfte liegen mittlerweile differenzierte Qualifikationsprofile vor (vgl. Wegweiser Sprachliche Bildung, 2011, S. 71–92), ebenso war es eine der Fragestellungen der Evaluation des Programms „sag mal was“ (Hasselbach, Schakib-Ekbatan, Roos & Schöler, 2007), ob die notwendigen Qualifikationen zur Diagnostik und Sprachförderung vorliegen.

Angesichts der Heterogenität der eingesetzten Verfahren in den Bundesländern sind zudem für die Anwendung eines Beobachtungsbogens andere Qualifikationsschwerpunkte erforderlich als bei Einsatz eines Testverfahrens oder Screenings – bei einem Beobachtungsverfahren ist nicht die Durchführung, sondern das Beobachtungs- und Bewertungsverhalten der pädagogischen Fachkraft Schwerpunkt der notwendigen Qualifikation.

Zur Durchführung, Auswertung, Interpretation und Mitteilung der Ergebnisse des Sprachstandsverfahrens dürfen nur Personen zugelassen werden, die hierzu qualifiziert sind.

Die durchführende Person⁹ ...

a) kennt die dem Sprachstandsverfahren zugrundeliegende Spracherwerbstheorie und kann diese kritisch reflektieren;

b) kennt die Stärken und Schwächen des verwendeten Verfahrens sowie von Sprachstandsverfahren allgemein, insbesondere im Hinblick auf die zu untersuchenden Zielgruppen;

c) kennt die Durchführung und kann auch auf Besonderheiten und Störungen flexibel reagieren;

d) kennt die Auswertungsanweisungen und kann Besonderheiten einer Diagnostik kompetent und fachlich angemessen berücksichtigen;

e) kennt die Interpretationsregeln und weicht von diesen nur in begründeten Fällen ab;

f) verfügt über die Kompetenz, eine angemessen differenzierte Ergebnismeldung an die Beteiligten/Betroffenen zu geben.

QUALITÄTSMERKMALE QUALIFIZIERUNG DER PÄDAGOGISCHEN FACHKRÄFTE

Für eine fachgerechte Durchführung muss ein Sprachstandsverfahren folgendes Kriterium erfüllen:

	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Die durchführende Person weist eine Qualifikation für die Durchführung von Sprachstandsverfahren auf.	Keine Angaben, keine vorgeschalteten Schulungen mit Abschlusstest / Evaluation der Kenntnisse der Teilnehmenden.	Die durchführende Person beherrscht mind. die 3 Aspekte c, d und e kompetent bzgl. des verwendeten Verfahrens.	Die durchführende Person beherrscht mind. die 4 Aspekte c, d, e und f kompetent bzgl. des verwendeten Verfahrens.	Die durchführende Person beherrscht alle o.g. Aspekte kompetent bzgl. des verwendeten Verfahrens.

Tabelle 11

ZEITLICHE ANFORDERUNGEN – DURCHFÜHRUNG, AUSWERTUNG & ERGEBNISRÜCKMELDUNG

Das Verfahren berücksichtigt die Konzentrationsfähigkeit des Kindes und die zeitliche Machbarkeit im pädagogischen Alltag.

Zwei Perspektiven sind bei der Betrachtung der zeitlichen Anforderungen an ein Sprachstandsverfahren zu beachten:

1. Vier- bis fünfjährige Kinder haben eine Konzentrationsspanne von 30 Minuten. Dieses Zeitfenster sollte bei

der Durchführung der Verfahren also nicht überschritten werden, da ansonsten mit einem erheblichen Einfluss auf die „Leistungsfähigkeit“ zu rechnen ist. Einzelne Verfahren tragen diesem Umstand Rechnung, indem entweder nur ausgewählte Bereiche geprüft werden sollen oder der Test in zwei Sitzungen durchgeführt werden kann.

2. Für die Person, die die Diagnostik durchführt, erscheint ein Aufwand für die zusätzlich durchzuführende Auswertung von max. 20 Minuten pro Kind als maximal zumutbar. Zusätzlich ist die Ergebnismeldung institutionsintern und an die Eltern mit ca. zehn bis 20 Minuten zu veranschlagen. Entsprechend sollte die Gesamtzeit für Durchführung, Auswertung und Ergebnismeldung etwa eine Stunde pro Kind betragen.

QUALITÄTSMERKMALE ZEITLICHE ANFORDERUNGEN

Für eine angemessene zeitliche Durchführbarkeit muss ein Sprachstandsverfahren folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Perspektive Kinder: zeitliche Dauer der Durchführung	Mehr als 30 Minuten; keine Angaben	Maximal 30 Minuten	Maximal 20 Minuten	Maximal 20 Minuten, abwechslungsreich gestaltet
Perspektive diagnostischer Aufwand: zeitliche Dauer der Durchführung, Auswertung & Dokumentation	Mehr als 40 Minuten, keine Angaben	Maximal 40 Minuten	Maximal 30 Minuten	Maximal 20 Minuten

Tabelle 12

MEHRSPRACHIGKEIT

Das Verfahren erfasst die besonderen Rahmenbedingungen von Kindern, die Deutsch als Zweitsprache erlernen. Nicht jedes mehrsprachige Kind hat per se einen Sprachförderbedarf.

Für Kinder, die Deutsch als Zweitsprache (DaZ) erlernen, müssen weitere Anforderungen berücksichtigt werden.^{10, 11} Bei Tests wird üblicherweise die Altersgruppe als Referenzgruppe für die Normierung ausgewählt; Gleichaltrige werden mit Gleichaltrigen im Hinblick auf ihr Sprachvermögen verglichen. Für mehrsprachige Kinder im Alter von vier bis fünf Jahren bedeutet das, ihre Kompetenzen in der Verkehrssprache Deutsch mit denen einsprachiger Kinder im Alter von vier bis sechs Jahren zu vergleichen. Der Sprachförderbedarf mehrsprachiger Kinder kann jedoch nicht ohne Weiteres an altersnormierten Erwerbsphasen einsprachiger Kinder gemessen werden, da u. a. Zeitpunkt des Erwerbs und die Kontakthäufigkeit und -intensität mit der Zweitsprache Deutsch erheblich variieren (Füssenich 2011). Bei Kindern mit fehlenden Deutschkenntnissen aufgrund einer anderen Erstsprache muss diese (sowie ggf. bestehende kulturelle Unterschiede) bei der Diagnostik, evtl. auch bei der Förderung, berücksichtigt werden. Das muss Stufe 1 jedes Auswahlverfahrens differenzieren können. Ein weiteres Kriterium ist die Testfairness: Kinder, die Deutsch als Zweitsprache erlernen, dürfen keine schlechteren Ergebnisse erzielen, nur weil ihnen kultu-

relles Wissen fehlt. Gibt man Kindern Aufgaben wie z. B. „Beschreibe, wie Du einen Schneemann im Winter bauen kannst“, so setzt dies zur Lösung v. a. kulturspezifisches Weltwissen („Was ist ein Schneemann?“) voraus. Erst darauf aufbauend kann es sprachliche Kompetenzen zeigen. Testfairness lässt sich über entsprechende (logistische) Analysen erfassen, indem getrennte Aufgabenanalysen für die Zielgruppen berechnet und die Itemparameter auf Unterschiedlichkeit getestet werden (sog. Differential Item Functioning). Dabei ist dann wichtig, dass Trennschärfe und Schwierigkeit der Aufgaben für die Gruppen identisch sind. Aufgaben, die unterschiedliche Werte aufweisen, benachteiligen bestimmte Zielgruppen.

Bei Kindern, die Deutsch als Zweitsprache lernen, ist es zudem wichtig, Informationen über ihr Umfeld und ihre Sprachbiographie zu erfassen. Hierbei ist darauf zu achten, dass die Fragen zu den Kontextbedingungen relevant für die Einschätzung des Sprachstandes sind, also empirisch-fachlich legitimiert werden können. So ist z. B. fraglich, ob der Erwerbsstatus der Eltern eine Relevanz für die Sprachbiografie des Kindes aufweist, während Fragen zur Literalität in der Familie durch den derzeitigen Forschungsstand als gut begründbar erscheinen.

Über die Fragen zur Sprachbiografie hinaus müssen für die sprachlichen Kompetenzen von Kindern mit Deutsch als Zweitsprache Normwerte vorliegen, nach Möglichkeit unter Berücksichtigung der Kontaktdauer (vgl. Schulz & Tracy 2012). Mit Blick auf die Diagnostik des Förderbedarfs muss das Verfahren eine Differenzierung ermöglichen, ob Mehrsprachigkeit vorliegt oder nicht.

Begründbare Kontextfragen zur angemessenen Erfassung der Sprachbiographie sind:

Welche Erst- bzw. Familiensprache(n) weist das Kind auf?

Wie ist die bisherige Sprachbiografie, und wieviele Kontaktmonate hatte das Kind bislang mit der Zielsprache? Wie viele (und vielfältige) Kontaktmonate mit der Zweitsprache gab es an welchen Aneignungsorten, welche Qualität und Quantität des zweitsprachlichen Inputs weist das Kind auf?

Werden die Erstsprachen der Kinder sowie mehrsprachige Sprachgebrauchsformen berücksichtigt (Ehlich 2007; Reich 2007; Schröder/Stölting 2005)?

Wie schätzen die Eltern das Sprachverhalten des Kindes und die sprachlichen Fähigkeiten in Erst- bzw. Familien- und Zweitsprache ein (Brizić 2007)? Wie ist die literarische Sozialisation bisher im Elternhaus verlaufen, wird es hierzu angeregt?

QUALITÄTSMERKMALE MEHRSPRACHIGKEIT

Für die angemessene Berücksichtigung mehrsprachiger Kinder müssen Sprachstandsverfahren folgende Kriterien erfüllen:

	Kriterien			
	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Werden die Erstsprache des Kindes sowie mehrsprachige Sprachgebrauchsformen berücksichtigt?	Keine Angaben; nicht berücksichtigt	Wird durch Abfrage, ob Deutsch L1 oder L2 ist, berücksichtigt.	Abfrage, ob Deutsch L1 oder L2 ist. Mind. 5-fach abgestuft (immer-nie) werden mehrsprachige sprl. Handlungsmuster abgefragt.	Es werden die Sprachen des Kindes und seine Sprachgebrauchsformen abgefragt, und es ist klar geregelt, wie diese in die Diagnose einfließen.
Wird die Sprachbiografie, insbesondere Kontaktmonate mit der Zweitsprache und Aneignungsorte und -qualität berücksichtigt?	Keine Angaben; keine Abfrage	Anzahl der Kontaktmonate wird abgefragt und berücksichtigt.	Anzahl der Kontaktmonate wird abgefragt und berücksichtigt. Aneignungsorte werden abgefragt.	Anzahl der Kontaktmonate wird abgefragt und berücksichtigt. Aneignungsorte werden abgefragt. Es werden klare Angaben gemacht, wie diese in die Diagnose einfließen.
Werden die Eltern gefragt, wie sie die sprachlichen Fähigkeiten ihres Kindes in der Erst- bzw. Familien- und Zweitsprache einschätzen?	Keine Angaben; keine Abfrage	Es wird die Information abgefragt, ob die Eltern nur, überwiegend oder (gar) nicht in L1 bzw. L2 mit dem Kind sprechen.	Zusätzlich erfolgt eine Einschätzung durch die Eltern, wie gut das Kind die L1 beherrscht.	Die Einschätzung der Eltern wird validiert.
Kann die Diagnose unterschiedlich ausfallen, je nachdem, ob Mehrsprachigkeit vorliegt oder nicht?	Keine Angaben; keine Berücksichtigung	Hinweis im Manual, ob Diagnose bei Kindern mit DaZ zulässig ist.	Differenzierung der Diagnose in: (1) Mangelnde Sprachbeherrschung Deutsch (L2) trotz ausreichender Beherrschung von L1 (2) Mangelnde Sprachbeherrschung Deutsch (L2) und L1 (3) Mangelnde Sprachbeherrschung Deutsch (L1) (4) Weder noch, d. h. ausreichende Sprachbeherrschung	
Separate Normwerte für DaZ-Kinder, nach Möglichkeit unter Berücksichtigung der Kontaktdauer.	Liegen nicht vor; unbekannt	An kleiner Stichprobe ermittelt (n < 100)	An ausreichender Stichprobe ermittelt (n > 100)	An ausreichender Stichprobe ermittelt, gegliedert nach Kontaktmonaten

Tabelle 13

SPEZIFITÄT DER DIAGNOSTIK

Das Verfahren gibt Hinweise darauf, welche Sprachfördermaßnahmen für das Kind in Frage kommen.

Der Fokus dieser Qualitätsmerkmale liegt auf der Auswahlgüte der Sprachverfahren. Die Spezifität der Diagnostik bezieht sich jedoch darauf, wie genau ein Ver-

fahren Aufschluss darüber gibt, welchen spezifischen Sprachförderbedarf ein Kind hat und welche konkreten Förderkonzepte infrage kommen. Dies ist auch notwendig, um eine Differenzierung zwischen Kindern mit Sprachförderbedarf und solchen mit einem sprachtherapeutischen Bedarf vornehmen zu können.¹² Entsprechend sollten die Ergebnisse des Verfahrens eine Unterscheidung in angemessene Sprachaneignung, (allgemeinen) Sprachförderbedarf und Sprachentwicklungsstörung ermöglichen.

Sprachförderung vs. Sprachtherapie

Es muss zwischen Kindern mit Sprachförderbedarf, Sprachauffälligkeit und Spacherwerbsstörung unterschieden werden. Ein Sprachförderbedarf liegt vor, wenn Kinder etwa drei bis vier Monate in der Entwicklung zurückgeblieben sind. Dies ist mit einer gezielten Förderung ausgleichbar. Bei sprachauffälligen Kindern, die etwas sechs Mona-

te in der Entwicklung zurückgeblieben sind, sind eine logopädische oder sprachtherapeutische Fachkraft und eine gezielte Therapie notwendig. Kinder, die eine Sprachtherapie benötigen, haben zudem meist nicht nur zeitliche Verzögerungen: therapiebedürftige ein- oder mehrsprachige Kinder zeigen eine „unausbalancierte“ Entwicklung, d. h., sie durchlaufen in einigen Bereichen nicht die typischen Erwerbsstufen.

Eine Diagnostik dient in der Pädagogik immer dem Zweck einer zielgerichteten Förderempfehlung (Ingenkamp & Lissmann, 2008). Diese ist umso effektiver, je spezifischer die Förderempfehlungen sind. Eine Diagnose „Risiko“ ist damit das unspezifischste Resultat einer Untersuchung, die im Rahmen z. B. der U8-Untersuchung als Screening-Ergebnis im ersten Schritt befriedigen mag (Kiese-Himmel & Rosenfeld 2012). Damit können Verfahren, die differenzierte Aussagen ermöglichen, diese Stärke auch in den hier vorgestellten Bewertungsrahmen einbringen. Gleichzeitig werden bereits Screening-

Instrumente, die „nur“ die Auswahl förderbedürftiger Kindern ermöglichen, als „ausreichend“ bewertet – wenn sie klare Angaben enthalten, wie eine Diagnostik des Förderbedarfs für die förderbedürftigen Kinder umzusetzen ist. Eine nach sprachlichem Bereich und kommunikativen Situationen differenzierende Diagnose, die auf für die identifizierten förderbedürftigen Bereiche spezifische und konkrete Förderempfehlungen verweist, ermöglicht dagegen eine sehr viel genauere Förderung.

QUALITÄTSMERKMALE SPEZIFITÄT DER DIAGNOSTIK

Für eine spezifische und differenzierte Diagnose müssen Sprachstandsverfahren folgendes Kriterium erfüllen:

	Nicht ausreichend	Ausreichend	Gut	Sehr gut
Mit dem Sprachstandsverfahren mögliche Diagnosekategorien	Keine Angaben; Risiko Ja / Nein (dichotom) ohne weiteren Hinweis, wie eine differenziertere Diagnose ermöglicht wird.	Risiko Ja / Nein (dichotom) sowie Verweis auf eine zweite Stufe, in der spezifisch geprüft wird.	(1) Differenzierung der Diagnose hinsichtlich mind. 3 sprachlicher Basisqualifikationen sowie (2) Verweis auf eine zweite Stufe, in der spezifisch geprüft wird	(1) Differenzierung der Diagnose hinsichtlich mind. drei sprachlicher Basisqualifikationen (2) Hinweise auf spezifische Fördermaßnahmen (3) Differenzierung der Diagnose, ob Förderung oder Therapie nötig ist

Tabelle 14

1 Die literale Basisqualifikation II, die den Schriftspracherwerb umfasst, ist für den Elementarbereich nicht wesentlich.

2 Auch für den seit 2011 vorliegenden Vergleich der Sprachstandsverfahren im Elementarbereich von Lisker (2011) wurde der Qualifikationsfächer nach Ehlich (2007) verwendet, dies allerdings in dichotomer Ausprägung und ohne Differenzierung nach der Qualität der Umsetzung: Wird in einem Verfahren nach der Wortschatzentwicklung gefragt, so wurde der Bereich „Semantische Basisqualifikation“ als „erfüllt“ klassifiziert.

3 Vgl. z. B. Heppt, Dragon, Berendes, Stanat & Weinert 2012

4 Vgl. Füssenich, 2011a: Vom Sprechen zur Schrift. Was Erwachsene über den Erwerb der Schrift im Elementarbereich wissen sollten; Füssenich 2011b: Glossar zur sprachlichen Bildung; sowie Ulich 2003a

5 Eine ausführliche Begründung der hier gewählten Grenzen für Cohens Kappa findet sich bei Landis und Koch (1977) wie auch Wirtz und Caspar (2002).

6 Der Determinationskoeffizient r^2 ist definiert als das Quadrat des Korrelationskoeffizienten. Mit dem Zahlenwert $r^2 \cdot 100$ wird der gemeinsame prozentuale Varianzanteil zweier Variablen ausgedrückt. Anders formuliert: Der Determinationskoeffizient gibt an, in welchem prozentualen Ausmaß die Varianz der einen Variablen durch die Varianz der anderen Variablen erklärt wird. Dieses Maß spiegelt wider, wie viel die zwei Variablen „gemeinsam“ haben.

7 Es sind in bedeutsamem Maße Variablen bekannt (z. B. Versuchsleitereffekte), die die Testleistung in unkontrollierter Weise beeinflussen (vgl. z. B. Rosenthal & Rosnow 1969). Sie können die interne Validität gefährden und zu Artefakten führen (Sarris & Reiß 2005).

8 Es ist nicht möglich, alle Kennwerte unabhängig voneinander zu optimieren. Insbesondere sind die Sensitivität und die Spezifität negativ miteinander korreliert: Wenn alle Kinder als „förderbedürftig“ klassifiziert werden, wird kein förderbedürftiges Kind übersehen (Sensitivität = 100 %). Analog gilt dies für die Spezifität.

9 Vgl. JC-Standards, Part II, Standardgruppe 6, APA, 1999

10 Dies kann bereits abgeleitet werden aus den Anforderungen, die aus psychometrischer Sicht bzgl. der aufgaben- und normierungsbezogenen Angemessenheit notwendig sind.

11 Ausführlich hierzu die Darstellungen von Jeuk (2009) und Lengyel (2012)

12 So existieren Untersuchungen für den deutschsprachigen Raum, denen zufolge drei bis 15 Prozent aller Kinder nicht eine Sprachförderung, sondern eine Sprachtherapie benötigen (Keilmann, Moein & Schöler 2012; Grimm 2011). Eine Förderung (statt Therapie) dieser Kinder ist nicht sinnvoll bzw. wird mit hoher Wahrscheinlichkeit keine Wirkungen aufweisen.

HANDLUNGSEMPFEHLUNGEN

Sprachstandsverfahren müssen so angelegt sein, dass sie **jedem Kind**, unabhängig vom Wohnort und dem sozialen Hintergrund, **Zugang zu einer bedarfsgerechten Sprachförderung** ermöglichen. Diese Chancengerechtigkeit ist durch die aktuelle Praxis in den Bundesländern nicht gegeben. Mit diesem Bewertungsrahmen liegen für zehn Handlungsfelder erstmals konkrete Qualitätsmerkmale vor, die auf einem interdisziplinären, wissenschaftlich fundierten Konsens fußen und, ohne der praktischen Reflexion vorzugreifen, die zentralen Anforderungen an pädagogische Fachkräfte sowie die Alltagstauglichkeit in den Blick nehmen. Sie umfassen

die sprachlichen Schlüsselkompetenzen, die für einen guten Start in die Schullaufbahn und die Teilhabe an der Gesellschaft von zentraler Bedeutung sind;

fundierte Hinweise darauf, wie Sprachstandsverfahren Kindern gerecht werden können, die Deutsch als Zweitsprache erlernen, ohne Mehrsprachigkeit zu stigmatisieren;

empirische Gütemerkmale, die für eine kontinuierliche Evaluation und Optimierung der Sprachstandsverfahren unabdingbar sind;

die Qualifizierung der pädagogischen Fachkräfte, die für die Sprachdiagnostik zuständig sind;

Orientierungswerte, wie sich die wissenschaftlich-pädagogischen Anforderungen an eine wirkungsvolle Diagnostik mit dem pädagogischen Alltag in Einklang bringen lassen.

Um die derzeitige Praxis nachhaltig zu verbessern und Kinder so früh wie möglich optimal zu fördern, ist ein intensiver Austausch zwischen Wissenschaft, Bildungspolitik und Praxis notwendig. Daher lädt das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache die verantwortlichen Akteure ein, ihre Verfahren kritisch zu prüfen und in einen Dialog darüber einzutreten, wie die derzeitige Praxis verbessert werden kann. Die Qualitätsmerkmale für Sprachstandsverfahren bilden einen Ausgangspunkt für diesen Austausch.

Die folgenden Handlungsempfehlungen geben Impulse für die Gestaltung einer solchen Schnittstelle zwischen

Wissenschaft, Bildungspolitik und Praxis.

Die Bundesländer haben auf unterschiedlichste Weise diagnostische Verfahren entwickelt oder implementiert. Dies ist legitim, darf aber nicht dazu führen, dass eine Beliebigkeit entsteht. Für eine wissenschaftlich legitimierte Auswahl diagnostik muss eine Reihe von Qualitätsmerkmalen beachtet werden. **Eine Vergleichbarkeit** – bei allen legitimen Unterschieden in der jeweiligen Ausgestaltung – zwischen den gewählten diagnostischen Verfahren **muss herstellbar sein**.

Sprachstandsverfahren sollten die zentralen Bereiche der sprachlichen Entwicklung eines Kindes sowie seiner (mehrsprachigen) Sprachbiografie erfassen. Gleichzeitig müssen sie wissenschaftliche Qualitätsmerkmale berücksichtigen und im pädagogischen Alltag praktikabel sein. Um diesen vielfachen Anforderungen gerecht zu werden, empfiehlt das Mercator-Institut, **ein interdisziplinäres Gremium ins Leben zu rufen**. Neben Vertreterinnen und Vertretern aus den betreffenden wissenschaftlichen Disziplinen sollten ebenso Akteure aus der Testentwicklung, der pädagogischen Praxis und der Bildungsadministration beteiligt sein.

Die stark variierenden Förderquoten weisen darauf hin, dass die Bundesländer bei der Definition eines Sprachförderbedarfs stark voneinander abweichen. Damit die Chance auf Förderung nicht vom Bundesland abhängt, empfiehlt das Mercator-Institut **eine zentrale Koordinierungsstelle zur einheitlichen Definition, ab wann Sprachförderung nötig ist**. Ein solches Gremium kann beispielsweise bei den Fachministerkonferenzen der Länder oder in der Bund-Länder-Initiative „Bildung durch Sprache und Schrift“ angesiedelt sein und auf Basis der Empfehlungen des Expertengremiums Hinweise zur Qualitätsverbesserung der landesspezifisch eingesetzten Verfahren geben.

Die **Anbindung an die wissenschaftliche Fachöffentlichkeit** muss gesichert sein. Nur so sind Fehlentwicklungen bei der Entwicklung und Umsetzung korrigierbar. Zudem benötigt die Wissenschaft die Erfahrung aus der Praxis, um den Forschungsstand systematisch zu verbessern. Dieser Austausch sollte z. B. über eine **jährlich stattfindende Fachkonferenz** unterstützt werden. Ebenso kann der Diskurs in wissenschaftlich einschlägigen Fachzeitschriften die einzelnen Entwicklungen dokumentieren und einer interdisziplinären Fachöffentlichkeit zugänglich machen.

Pädagogische Fachkräfte müssen nicht nur für die Durchführung der Sprachstandsverfahren geschult werden, sondern benötigen für eine fundierte Diagnostik ebenso allgemeine Kenntnisse über den kindlichen Spracherwerb. **Pädagogische Fachkräfte sollten daher gezielt auf die Durchführung der Sprachstandsverfahren vorbereitet werden**. Sowohl in der beruflichen wie auch akademischen Ausbildung braucht es Module, die angehende Fachkräfte, entsprechend ihres Profils, auf Sprachdiagnostik und Sprachförderung vorbereiten.

Neben diesen mittelfristigen Gestaltungsmöglichkeiten wird es auf lange Sicht nötig sein, die eingesetzten **Verfahren auf ihre Wirksamkeit zu überprüfen**. In einer Querschnittsstudie müssten dafür zunächst die einzelnen Verfahren unter gleichen Einsatz- und Rahmenbedingungen auf Basis der hier erarbeiteten Qualitätsmerkmale überprüft werden. Mithilfe einer Längsschnittstudie ließe sich dann überprüfen, ob die Kinder, die im Alter von vier Jahren getestet und ggf. Sprachförderung erhalten haben, am Ende des ersten Schuljahres über die für dieses Alter definierten sprachlichen Kompetenzen verfügen.

KONTINUIERLICHE VERBESSERUNG DER EINGESETZTEN VERFAHREN DURCH FORMATIVE EVALUATION



Abbildung 2: Mögliches Vorgehen zur Etablierung und kontinuierlichen Verbesserung eines Verfahrens

Prof. Dr. Michael Becker-Mrotzek ist Direktor des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache. Er gehört zur Autorengruppe der Expertise „Bildung durch Sprache und Schrift“, Grundlage der Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung. Michael Becker-Mrotzek ist Mitglied im Beirat des Zentrums für Mehrsprachigkeit und Integration Köln (ZMI) und stellvertretender wissenschaftlicher Leiter des Zentrums für LehrerInnenbildung der Universität zu Köln.

Prof. Dr. Dr. h. c. Konrad Ehlich (Honorarprofessor an der Freien Universität Berlin) hat von 1992 bis 2007 das Institut für Deutsch als Fremdsprache der Ludwig-Maximilians-Universität München geleitet. 2007 hat er dem Bundesministerium für Bildung und Forschung einen Bericht über Anforderungen an Sprachstandsverfahren im Elementarbereich vorgelegt. 2012 war er an der Erstellung einer Expertise für erfolgreiche Sprachförderung im Auftrag des Berliner Senats beteiligt. Konrad Ehlich ist Mitglied des wissenschaftlichen Beirats des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache sowie Mitglied der Bewertungskommission für Integrationskurse des Bundesinnenministeriums.

Prof. Dr. phil. Iris Füssenich hat eine Professur für den Förderungsschwerpunkt Sprache und Kommunikation an der sonderpädagogischen Fakultät der Pädagogischen Hochschule Ludwigsburg inne. In Baden-Württemberg hat sie die Umsetzung des Orientierungsplans Bildung und Erziehung mit dem Schwerpunkt Frühförderung begleitet. Sie ist Mitglied im wissenschaftlichen Beirat der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WIFF).

Prof. Dr. Hartmut Günther ist emeritierter Professor für Deutsche Sprache und Literatur und ihre Didaktik an der Universität zu Köln. Zu seinen Forschungsschwerpunkten zählen Schriftspracherwerb und Rechtschreibung. Er hat zahlreiche Schulbücher und Arbeitsmaterialien für den Deutschunterricht in der Grundschule herausgegeben.

Prof. Dr. Marcus Hasselhorn ist geschäftsführender Direktor des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF). Er gehört zu den Autoren der Expertise „Bildung durch Sprache und Schrift“, Grundlage der Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung. Er ist Vorstandsmitglied der 2013 gegründeten Gesellschaft für empirische Bildungsforschung. Von 2006 bis 2008 war er Präsident der Deutschen Gesellschaft für Psychologie. Marcus Hasselhorn ist Mitglied des wissenschaftlichen Beirats des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache.

Dr. Michaela Hopf leitet die Fachgruppe „Pädagogische Konzepte für die frühe Kindheit“ im Deutschen Jugendinstitut (DJI). Einer ihrer Arbeitsschwerpunkte ist sprachliche Bildung und Förderung in der frühen Kindheit. Das DJI setzt u. a. die Qualifizierungsoffensive „Sprachliche Bildung und Förderung für

Kinder unter Drei“ sowie die Weiterbildungsinitiative Frühpädagogische Fachkräfte (WIFF) um.

apl. Prof. Dr. Stefan Jeuk ist akademischer Direktor der Abteilung Deutsch an der Pädagogischen Hochschule Ludwigsburg. Seine Arbeitsschwerpunkte sind u. a. Spracherwerb und Mehrsprachigkeit sowie Sprachförderung und Förderunterricht bei Kindern mit Migrationshintergrund. Er hat mehrere Jahre als Lehrer an einer Förderschule in Ludwigsburg gearbeitet.

Prof. Dr. Drorit Lengyel ist Professorin für Erziehungswissenschaft an der Universität Hamburg. Zuvor hatte sie eine Juniorprofessur für Sprachdiagnostik im Kontext von Zwei- und Mehrsprachigkeit inne. Darüber hinaus war sie im Modellprogramm von Bund und Ländern „Förderung von Kindern und Jugendlichen mit Migrationshintergrund“ (FörMig) tätig.

Dr. Uwe Neugebauer ist wissenschaftlicher Mitarbeiter am Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache. Neben der psychologischen Diagnostik und Beratung nach Banküberfällen liegt sein Schwerpunkt in der Evaluation von sprachlichen Fördermaßnahmen, so u. a. im Rahmen des BLK-Programms FörMig oder auch des Programms „sag mal was“ der Baden-Württemberg-Stiftung.

Prof. Dr. Argyro Panagiotopoulou ist Professorin für Bildung und Entwicklung in Früher Kindheit an der Universität zu Köln. Zu ihren Forschungsschwerpunkten gehören Sprachentwicklung und Sprachförderung im Elementar- und Primarbereich sowie Inklusion und Umgang mit Heterogenität. Argyro Panagiotopoulou ist Mitglied des wissenschaftlichen Beirats des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache.

Prof. Dr. Petra Stanat ist Direktorin des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin. Das IQB führt im Auftrag des Bundes und der Länder regelmäßig Ländervergleichsstudien durch, die untersuchen, inwieweit Schülerinnen und Schüler die national verbindlichen Bildungsstandards erreichen. Sie gehört zu den Autoren der Expertise „Bildung durch Sprache und Schrift“, Grundlage der Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung. Seit 2012 ist sie Mitglied im Integrationsbeirat der Bundesregierung. Petra Stanat ist Mitglied des wissenschaftlichen Beirats des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache.

Dr. habil. Jürgen Wilbert arbeitet im Department Heilpädagogik und Rehabilitation der Universität zu Köln. Er ist Mitherausgeber der Zeitschrift „Empirische Sonderpädagogik“, im Leitungsgremium des Zentrums für empirische Inklusionsforschung bei Lern- und Entwicklungsbarrieren (ZEIF) tätig sowie im Editorial Board des Journals „Insights on Learning Disabilities“. Der Schwerpunkt von Dr. Jürgen Wilbert sind methodische Fragestellungen sowie komplexe Fragen der Abbildung von Lernverlaufsdiagnostik.

- Autorengruppe Bildungsberichterstattung (2012). Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf. Bielefeld: Bertelsmann Bartram, D. (2008). EFPA REVIEW MODEL FOR THE DESCRIPTION AND EVALUATION OF PSYCHOLOGICAL TESTS. TEST REVIEW FORM AND NOTES FOR REVIEWERS. Internes Dokument.
- Brizic, K. (2009). Bildungsgewinn bei Sprachverlust? Ein soziolinguistischer Versuch, Gegensätze zu überbrücken. In: Streitfall Zweisprachigkeit – The Bilingualism Controversy (pp. 133–143). VS Verlag für Sozialwissenschaften.
- Chlosta, C. & Ostermann, T. (2007) Warum fragt man nach der Herkunft, wenn man die Sprache meint?. In: Bildungsforschung Band 14, Migrationshintergrund von Kindern und Jugendlichen: Wege zur Weiterentwicklung der amtlichen Statistik, 55–66
- Die Bundesregierung (2007). Der Nationale Integrationsplan. Neue Wege – neue Chancen. Berlin. Ehlich, K. (2007). Sprachaneignung und deren Feststellung bei Kindern mit und ohne Migrationshintergrund: Was man weiß, was man braucht, was man erwarten kann. In: Ehlich, K. (Hrsg.): Anforderungen an Verfahren der regelmäßigen Sprachstandsfeststellung als Grundlage für die frühe und individuelle Förderung von Kindern mit und ohne Migrationshintergrund. Berlin: BMBF. S. 11–120.
- Ehlich, K., Bredel, U., Reich, H. H. (Hrsg.) (2008). Referenzrahmen zur altersspezifischen Sprachaneignung – Forschungsgrundlagen. Bildungsforschung Band 29. Berlin: BMBF.
- Evers, A. (2001). Beurteilungssystem für die Qualität von Tests. Commissie Testaangelegenheden Nederland van het Nederlands Instituut van Psychologen/ NIP. COTAN = Committee On Test Affairs Netherlands. Übersetzt durch PD Dr. Carmen Hagemeister, Technische Universität Dresden.
- Füssenich, I. (2011a): Vom Sprechen zur Schrift. Was Erwachsene über den Erwerb der Schrift im Elementarbereich wissen sollten. WiFF Expertisen 9. München
- Füssenich, I. (2011b). Glossar zur sprachlichen Bildung. Von Spracherwerbstheorien bis zur Sprachförderung – Eine Einführung. In: DJI – Impulse, Heft 4, S. 25–28.
- Grimm, H. (2012). Störungen der Sprachentwicklung: Grundlagen – Ursachen – Diagnose – Intervention – Prävention. Hogrefe: Göttingen. Grießhaber, W. (2005). Sprachstandsdiagnose im kindlichen Zweitspracherwerb: Funktional-pragmatische Fundierung der Profilanalyse. Veröffentlichungsort unbekannt. S. 1–58.
- Hacker, D. (2002). Phonologie. In: Baumgartner, Stephan/Füssenich, Iris (Hrsg.): Sprachtherapie mit Kindern. Grundlagen und Verfahren. München/Basel, 5. überarbeitete Auflage, S. 15–79.
- Hartmann, P. (2002). Indikator. Wörterbuch der Soziologie, Stuttgart: Lucius & Lucius, 223–224. Hartmann, E. (2002). Möglichkeiten und Grenzen einer präventiven Intervention zur phonologischen Bewusstheit von lautsprachgestörten Kindergartenkindern. Freiburg: Sprachimpuls.
- Hasselbach, P., Schakib-Ekbatan, K., Roos, J. & Schöler, H. (2007). Die Bewertung der Sprachfördermaßnahmen aus der Sicht der Förderkräfte – Interviews. Verfügbar unter: http://www.sagmalwas-bw.de/projekt01/media/pdf/EVAS_Bericht2.pdf [13.04.2013].
- Hasselhorn, M., Schneider, W., & Marx, H. (2000). Diagnostik von Lese- Rechtschreibschwierigkeiten. Tests und Trends, N. F. Band 1. Hogrefe-Verlag, Göttingen.
- Heilmann, B. (2012). Diagnostik & Förderung-leicht gemacht: Deutsch als Zweitsprache. Ein Praxishandbuch mit Video-DVD. Klett Ernst/Schulbuch.
- Heppert, B., Dragon, N., Berendes, K., Stanat, P. & Weintert, S. (2012). Beherrschung von Bildungssprache bei Kindern im Grundschulalter. Diskurs Kindheits- und Jugendforschung. Heft 3, S. 349–356.
- Ingenkamp, K.-H., & Lissmann, U. (2008). Lehrbuch der Pädagogischen Diagnostik. Beltz: Weinheim.
- Jansen, H., Mannhaupt, G., Marx, H. & Skowronek, H. (1999). Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC). Göttingen: Hogrefe.
- Jeuk, S. (2009). Aktuelle Verfahren zur Einschätzung des Stands der Sprachaneignung im Grundschulalter. In: Jeuk, Stefan; Schmid-Barkow, Ingrid (Hrsg.): Differenzen diagnostizierten und Kompetenzen fördern im Deutschunterricht. Freiburg: Fillibach, S. 61–81.
- Keilmann, A., Moein, G., & Schöler, H. (2012). Does the SETK 3–5 detect clinically diagnosed language impairment? HNO, 60(1), 63.
- Kany, W., & Schöler, H. (2010). Fokus Sprachdiagnostik: Leitfaden zur Sprachstandsbestimmung im Kindergarten (2. Auflage). Berlin: Cornelsen Scriptor.
- Kersting, M. (2006). Zur Beurteilung der Qualität von Tests: Resümee und Neubeginn. Psychologische Rundschau, 57 (4), S. 243–253
- Kiese-Himmel, C., & Rosenfeld, J. (2012). Analyse aktueller Untersuchungsinstrumente zur Früherkennung von Auffälligkeiten in Sprechen und Sprache in der pädiatrischen Vorsorgeuntersuchung U8. Das Gesundheitswesen, 74(10), S. 661–672.
- Klauer, K.C. (1987): Kriteriumsorientiertes Testen: Der Schluß auf den Itempool [Criterion-referenced testing: The inference to the item pool]. Zeitschrift für differentielle und diagnostische Psychologie, 8, S. 141–147.
- Kniffka, G., & Siebert-Ott, G. (2007). Deutsch als Zweitsprache. Paderborn: Schöningh Landis, J.R., Koch, G.G (1977). The measurement of observer agreement for categorical data. In: Biometrics. 33, 159–174.
- Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). The feasibility of universal screening for primary speech and language delay: Findings from a systematic review of the literature. Dev Med Child Neurol; 42: 190–200.
- Lengyel, D. (2012). Sprachstandsfeststellung bei mehrsprachigen Kindern im Elementarbereich. Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF) Expertise 29. Herausgeber: Deutsches Jugendinstitut e. V. (DJI), München.
- Lienert, G. A. & Raatz, U. (1998). Testaufbau und Testanalyse. Weinheim: Beltz
- Lisker, A. (2010). Sprachstandsfeststellung und Sprachförderung im Kindergarten und beim Übergang in die Schule. Expertise im Auftrag des Deutschen Jugendinstituts. München: Deutsches Jugendinstitut e.V.
- Lisker, A. (2011). Additive Maßnahmen zur vorschulischen Sprachförderung in den Bundesländern. Expertise im Auftrag des Deutschen Jugendinstituts.
- Marx, H. (1992). Methodische und inhaltliche Argumente für und wider eine frühe Identifikation und Prädiktion von Lese-Rechtschreibschwierigkeiten. Diagnostica, 38 (3), 249–268.
- Marx, P. & Lenhard, W. (2010). Diagnostische Merkmale von Screeningverfahren. In M. Hasselhorn & W. Schneider (Hrsg.), Frühprognose schulischer Kompetenzen. Göttingen: Hogrefe.
- Marx, P., & Weber, J. (2006). Vorschulische Vorhersage von Lese- und Rechtschreibschwierigkeiten. Neue Befunde zur prognostischen Validität des Bielefelder Screenings (BISC). Zeitschrift für Pädagogische Psychologie, 20 (4), 251–259
- Mayer, H. O. (2004). Interview und schriftliche Befragung: Entwicklung, Durchführung und Auswertung. Oldenbourg Wissenschaftsverlag, München.
- Moein, G., & Schöler, H. (2012). Werden mit dem SETK 3–5 klinisch diagnostizierte Sprachentwicklungsstörungen erfasst? HNO, 60(1), S. 63–71.
- Neugebauer, U. (2010). Keine Outcomes trotz Kompetenzüberzeugung? Qualifikationen und Selbsteinschätzungen von Sprachförderkräften in KiTa's. Empirische Sonderpädagogik (2), S. 34–47.
- Plante, E, Vance, R. (1994). Selection of preschool language tests: A data-based approach. Lang Speech Hear Serv in Schools; 25: 15–24 .
- Reich, H. H. (2006). Tests und Sprachstandsmessungen bei Schülern und Schülerinnen, die Deutsch nicht als Muttersprache haben. In: Bredel, U. et al.: Didaktik der deutschen Sprache. Bd. 2. 2. Auflage. Paderborn, UTB- Schöningh. S. 914–923.
- Reich, H. H. (2007). Forschungsstand und Desideratenaufweis zu Migrationslinguistik und Migrationspädagogik für die Zwecke des „Anforderungsrahmen“. In: Ehlich, K. (Hrsg.): Anforderungen an Verfahren der regelmäßigen Sprachstandsfeststellung als Grundlage für die frühe und individuelle Förderung von Kindern mit und ohne Migrationshintergrund. Berlin: BMBF. S. 121–170.
- Reich, H. H. (2008). Die Sprachaneignung von Kindern in Situationen der Zweid und Mehrsprachigkeit. In: Ehlich, K.; Bredel, U.; Reich, H. H. (Hrsg.): Referenzrahmen zur altersspezifischen Sprachaneignung – Forschungsgrundlagen. Berlin: BMBF. S. 163–169.
- Rosenthal, R., & Rosnow, R. L. (1969). The volunteer subject. In R. Rosenthal & R. L. Rosnow (Eds.), Artifact in behavioral research (pp. 59–118). New York: Academic Press. Saris & Reib, 2005.
- Schöler, H. & Roos, J. (2011). Die Ergebnisse des Projekts EVAS, der Evaluationsstudie zur Sprachförderung von Vorschulkindern in Heidelberger und Mannheimer Kindergärten. In: Baden-Württemberg Stiftung, Sag mal was – Sprachförderung für Vorschulkindern. Tübingen: Francke Verlag.
- Schröder, C., & Stöltig, W. (2005). Mehrsprach orientierte Sprachstandsfeststellungen für Kinder mit Migrationshintergrund. In: Gogolin, I., Neumann, U., Roth, H.-J. (Hrsg.): Sprachdiagnostik bei Kindern und Jugendlichen mit Migrationshintergrund. Dokumentation einer Fachtagung am 14. Juli 2004 in Hamburg. Münster: Waxmann 2005, S. 59–74.
- Schulz, P., & Tracy, R. (2011). LiSe-DaZ – Linguistische Sprachstandserhebung– Deutsch als Zweitsprache. Göttingen: Hogrefe Testbeurteilungssystem des Testkuratoriums (TBS– TK) der Föderation Deutscher Psychologenvereinigungen (2009)
- Tippelt, S., Kühn, P., Großheirich, N. & von Suchodoletz, W. (2011). Diagnostische Zuverlässigkeit von Sprachtests und Elterntesting bei Sprachentwicklungsstörungen. Laryngo-Rhino-Otol; 90: 421–427
- Tracy, R. (2007). Linguistische Grundlagen der Sprachförderung: Wieviel Theorie braucht (und verlangt) die Praxis? In: Ahrenholz, Bernt (Hrsg.). Deutsch als Zweitsprache. Voraussetzungen und Konzepte für die Förderung von Kindern und Jugendlichen mit Migrationshintergrund. Baltmannsweiler: Schneider Verlag Hohengehren. S. 17–29.
- Ulich, M. (2003a). Förderung von Kindern mit einem anderen kulturellen Hintergrund. Bundesministerium für Familien, Senioren, Frauen und Jugend (Hrsg.): Auf den Anfang kommt es an, 141–153.
- Ulich, M. (2003b). Literacy und sprachliche Bildung im Elementarbereich. KiTa spezial. Sonderausgabe von KiTa aktuell, S. 12.
- Valtin, R. (2012). Phonologische Bewusstheit: Ein kritischer Blick auf ein modisches Konstrukt. Frühe Bildung, 1(4), S. 223–225.
- Wirtz, M., & Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität. Hogrefe-Verlag, Göttingen.

Herausgeber

Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache
Innere Kanalstraße 15
50823 Köln
0221 – 470 5718
info@mercator.uni-koeln.de
www.mercator-institut-sprachfoerderung.de

**Über das Mercator-Institut für Sprachförderung
und Deutsch als Zweitsprache**

Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache ist ein von der Stiftung Mercator initiiertes und gefördertes Institut der Universität zu Köln. Ziel des Instituts ist es, langfristig die sprachliche Bildung an deutschen Schulen zu verbessern, damit alle Schülerinnen und Schüler gute Chancen auf eine erfolgreiche Bildungskarriere haben. Es berät Hochschulen dabei, Deutsch als Zweitsprache in der Lehrerbildung zu verankern, fördert, vermittelt und betreibt anwendungsorientierte Forschung und trägt zur Qualifizierung von Lehrenden in Schule und Hochschule bei. Darüber hinaus beobachtet es Bildungspraxis-, -verwaltung und -politik.



Stiftung
Mercator

**Verantwortlich**

Prof. Dr. Michael Becker Mrotzek

Wissenschaftliche Projektleitung

Dr. Uwe Neugebauer

Redaktion

Anna Kleiner

In das Dokument ist die Expertise aller Autoren und Autorinnen eingeflossen. Bestehende Texte einzelner Autoren und Autorinnen sind eingeflossen, ohne dass dies immer als Zitat gekennzeichnet wurde. Andere Quellen wurden dagegen immer explizit belegt. Diese Publikation darf, unter Einhaltung der gängigen Zitierregeln und mit Angabe der Quelle, gern weiterverwendet werden:

Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache (Hrsg.)
(2013): Qualitätsmerkmale für Sprachstandsverfahren im Elementarbereich.
Ein Bewertungsrahmen für fundierte Sprachdiagnostik in der Kita. Köln.

Lektorat

TextSchleiferei, Köln

Gestaltung

kikkerbillen – Büro für Gestaltung, Köln



