

Die Problematik der Evidenz von Testverfahren am Beispiel der Aussprachestörung

Annette Fox-Boyer*

Zusammenfassung

Evidenzbasiertes therapeutisches Handeln kann nur auf der Grundlage von Ergebnissen aus ebenfalls evidenzbasierten Diagnostikverfahren stattfinden. Bislang wird im deutschsprachigen Raum die Notwendigkeit der Überprüfung von sprachtherapeutischen Diagnostikverfahren im Hinblick auf ihre Evidenzbasierung kaum diskutiert. Im vorliegenden Beitrag soll am Beispiel der kindlichen Aussprachestörung dargestellt werden, welche Kriterien Diagnostikverfahren erfüllen müssen, um als evidenzbasiert gelten zu können. Des Weiteren werden aktuell verfügbare Verfahren kritisch betrachtet. Es zeigt sich, dass die Verfahren deutliche Mängel im Hinblick auf die geforderten Kriterien aufweisen, so dass eine Überprüfung der tatsächlichen Evidenz der Verfahren notwendig wird.

Schlüsselwörter: Evidenz, Testverfahren, Aussprachestörungen, Deutsch



Originalbeitrag
open access

* **Korrespondenz:**
Annette Fox-Boyer
a.fox@eufh.de

Zitation: Fox-Boyer, A. (2018) Die Problematik der Evidenz von Testverfahren am Beispiel der Aussprachestörung. Sprachtherapie aktuell: Forschung - Wissen - Transfer I: Schwerpunktthema: Von der Forschung in die Praxis: e2018-02; doi: 10.14620/stadbs181202

1. Einleitung

Sprachtherapeuten sind aufgefordert, ihre Patienten nach bestmöglicher Evidenzlage zu behandeln. Dies bedeutet, dass sie eine Behandlung der Ziele entsprechend der Bedürfnisse des Patienten anstreben sollten, die sowohl auf bestmöglicher externer (forschungsbasierter) als auch interner (therapeutische Expertise) Evidenz basiert (Dollaghan, 2007; siehe auch Hanne, dieser Band). Die Grundlage einer jeglichen Behandlung allerdings stellt die korrekte Identifikation eines behandlungsbedürftigen Problems dar. Diese korrekte Identifikation ist an evidenzbasierte Diagnostikinstrumente gebunden, was bedeutet, dass diese Instrumente verlässlich, genau und gerecht sein müssen (Limbrick et al. 2013). Diagnostische Instrumente stellen für Patienten den Zugang auf Behandlungsanspruch dar (Crais, 2011). Damit dies ausreichend gewährleistet werden kann, sind verschiedene Kriterien bei der Konstruktion eines Diagnostikinstrumentes zu beachten und es ist zu überprüfen, ob das Instrument tatsächlich sensitiv und spezifisch das zu messende Problem erfasst. Zu diesen Kriterien zählen die Gewährleistung von Validität, Reliabilität, Objektivität und eine vorliegende Normierung des Instruments. Interessanterweise wird im deutschsprachigen Raum, zumindest in Bezug auf Sprachentwicklungsstörungen, kaum diskutiert, inwieweit die existierenden diagnostischen Verfahren die genannten Kriterien erfüllen. Selbst die S2k-Leitlinie „Diagnostik von Sprachentwicklungsstörungen“ (De Langen-Müller et al 2011) verzichtet auf eine derartige Überprüfung. Lediglich Beushausen befasste sich bereits 2008 mit dieser Thematik. Erstmals fachspezifisch wurde eine Publikation zu diesem Thema im Hinblick auf kindliche Aussprachestörungen in diesem Jahr veröffentlicht (Fox-Boyer, Albrecht & Clausen, 2018). Dies steht im Gegensatz zu zahlreichen internationalen Publikationen, die das Thema theoretisch für die Sprachtherapie diskutieren (Dollaghan, 2004; Friberg, 2010, Messick, 1980, 1989; Peña et al, 2006), aber auch spezifisch für den Bereich der kindlichen Aussprachestörungen (Dollaghan 2004, Flipsen u. Ogiela 2015, Kirk u. Vigeland 2014).

2. Kindliche Aussprachestörungen

Kindliche Aussprachestörungen können mögliche Schwierigkeiten im Bereich Perzeption, Artikulation/motorische Produktion und/oder phonologischer Repräsentation der Sprechsegmente (Konsonanten und Vokale), der Phonotaktik (Silben und Wortformen) und der Prosodie (lexikalische und grammatikalische Töne, Rhythmus,,

Betonung und Intonation) betreffen, die sich negativ auf die Verständlichkeit und Akzeptanz der Kinder auswirken können (IEPMCS, 2012). Kinder mit Aussprachestörungen stellen daher eine extrem heterogene und große Gruppe dar (Prävalenz: ca. 16 % betroffener Vorschulkinder (3-6 Jahre) Fox-Boyer, 2014a). Je nach Subtyp tragen die Kinder ein ausgeprägtes Risiko für spätere Lese- Rechtschreibstörungen und zeigen schlechtere Chancen für eine gesunde psychosoziale und erfolgreiche schulische/berufliche Entwicklung (Felsenfeld et al 1994; Leitão und Fletcher, 2004; McCormack et al 2010; Schnitzler, 2015). Daher ist eine frühzeitige Identifikation und Behandlung der betroffenen Kinder von großer Bedeutung.

3. Kriterien zur Evidenz von Testverfahren für Aussprachestörungen

Im deutschsprachigen Raum werden in der Regel Bilderbenennverfahren als Kerninstrument benutzt, um kindliche Aussprachestörungen zu identifizieren. In der internationalen Diskussion findet sich eine gewisse Einigkeit darüber, dass dieses Vorgehen verlässliche Ergebnisse im Hinblick auf die Bewertung der Ausspracheentwicklung bietet, wenn festgelegte Kriterien erfüllt sind, die die Inhaltsvalidität gewährleisten. Zu diesen Kriterien zählen phonetisch/phonologische und lexikalische Aspekte. Ein Untersuchungsinstrument sollte ca. 100 Items beinhalten, alle Phoneme (Konsonanten und Vokale) sollten mindestens viermal pro Wortposition überprüft werden, die zentralen Konsonantenverbindungen, mögliche Betonungsmuster und Silbenstrukturen sollten enthalten sein. Zudem sollten Wörter unterschiedlicher Silbenlänge enthalten sein. Mit dem Verfahren sollte es möglich sein, die sprachspezifischen typischen und atypischen phonologischen Prozesse auf der Basis einer phonetischen Transkriptionen auszuwerten (z. B. Bernhardt und Holdgrafer, 2012; Eisenberg und Hitchcock, 2010; James, 2001; Kirk und Vigeland, 2014, 2015). Die ausgewählten Items müssen dem Wortschatz kleiner Kinder entsprechen und eindeutig abbildbar sein. Das Verfahren sollte normiert sein und in seiner Durchführungsdauer angemessen.

Neben der Inhaltsvalidität sollte insbesondere die diagnostische Validität überprüft worden sein, das bedeutet, es sollte nachgewiesen worden sein, dass eine spezifische und sensitive Erkennung von kindlichen Aussprachestörungen möglich ist. Des Weiteren sollte nachgewiesen werden, dass unterschiedliche Untersucher eine hohe Übereinstimmung der diagnostischen Ergebnisse zeigen (Beurteiler-Reliabilität) und dass das Testergebnis nicht von der Tages Performanz abhängig ist (Test-Retest-Reliabilität). Die Objektivität des Verfahrens sollte mithilfe eines Handbuchs, welches klare Informationen über die Qualifikation des Testers, die Durchführung, Auswertung und Interpretation des Tests bietet, nachgewiesen

sein.

4. Kritische Reflexion deutschsprachiger Verfahren

Für den deutschsprachigen Raum liegen zur Zeit sieben Verfahren vor, die eine phonologische Prozessanalyse ermöglichen und damit als aktuell zu werten sind (siehe Tabelle 1).

Alle Verfahren verfügen über ein Handbuch, wobei sich diese in ihrer Ausführlichkeit und Genauigkeit insbesondere in Bezug auf die Inhaltsvalidität die Auswertungs- und Interpretations-Objektivität deutlich unterscheiden. Die wenigsten Verfahren erläutern, nach welchen Kriterien die Testitems ausgewählt wurden, so dass ein Vergleich der phonetisch-phonologischen Kriterien schwierig ist.

Die Verfahren unterscheiden sich in der Itemanzahl, wobei SET (57) und APP (71) die wenigsten Items aufweisen und der AVAK (115) über den vorgeschlagenen 100 Items liegt. Die weiteren Verfahren weisen 82-96 Items auf. Damit lässt sich erklären, warum die geforderten Überprüfungen von Konsonanten, Vokalen und Konsonantenverbindungen nicht immer eingehalten werden können. Der Faktor Silbenstruktur wird von allen Verfahren beachtet, wobei der SET 3-5 keine Angaben dazu macht. Der Faktor mehrsilbige, mono-morphematische Wörter wird nur in zwei Verfahren genannt (PLAKSS-II/ APP) und nur zwei Verfahren verfügen über eine ausreichende Anzahl entsprechender Items (PLAKSS-II /PAP). Die Überprüfung unterschiedlicher Wortbetonungen wird in drei Verfahren erwähnt, allerdings enthalten die Verfahren bis auf die PLAKSS-II nur je fünf Items, die nicht trochäisch sind. Zwei Verfahren überprüfen die Wortbetonung anhand von zusätzlichen Testitems.

In der Regel gibt es keinerlei Hinweise auf Kriterien, die zur Auswahl des genutzten Wortschatzes angewendet wurden. Normen im Hinblick auf phonologische Prozesse stehen lediglich für die PLAKSS-II zur Verfügung. Der SET 3-5 bietet Normwerte auf der segmentellen Ebene des Phon- und Phonem-Erwerbs, aber nicht auf der Ebene der phonologischen Prozesse. Die Verfahren AVAK und PAP bieten keinerlei Altersvergleich, während die Verfahren PDSS, LOGO, APP auf unterschiedliche Normdaten, die mithilfe der PLAKSS-I erworben wurden, verweisen.

Die diagnostische Validität wurde bislang lediglich für die Verfahren PLAKSS-I/II (Fox & Dodd, 2001; Fox-Boyer und Meyer in Vorbereitung) und SET 3-5 überprüft. Im Hinblick auf die Reliabilität liegen zu den Verfahren AVAK; PDSS, LOGO und APP keine Angaben vor. Die PAP beschreibt, dass Untersuchungen zur Test-Retest-Reliabilität und Interrater Reliabilität hohe Übereinstimmungen zeigten hätten, bietet aber keine Belege dafür. Der SET 3-5 wies hinsichtlich der Test-Retest-Reliabilität eine hohe Korrelation für den Vergleich der Fehlbildungsrate der Phoneme zu den zwei Testzeitpunkten. Untersuchungen zur Test-Retest-Reliabilität und Interrater

Tabelle 1: Deutschsprachige Verfahren zur Überprüfung der kindlichen Aussprache

Test	Altersspanne
AVAK-Test Analyse zu Aussprachestörungen bei Kindern, (Hacker u. Wilgermein 2002)	k.A.*
PAP Pyrmonter Aussprache Prüfung, (Babbe 2003)	k.A.*
PDSS Patholinguistische Diagnostik bei Sprachentwicklungsstörungen –Anteil Phonetik/Phonologie, (Kauschke u. Siegmüller 2010)	2;0-6;11
LOGO-Ausspracheprüfung Logo-Ausspracheprüfung zur differenzierten Analyse von Dyslalien, (Wagner 2011)	3-7
PLAKSS-II Psycholinguistische Analyse kindlicher Aussprachestörungen, (Fox-Boyer 2014b)	2;6-7;11
APP Ausspracheprüfung für die Praxis - Österreich, (Brunner 2015)	3-6 Jahre aber auch früher oder später einsetzbar
SET 3-5 Sprachstandserhebungstest für Kinder im Alter zwischen 3 und 5 Jahren, (Petermann 2016)	3;0-5;11

*k.A.: keine Angabe

ter Reliabilität der PLAKSS-II zeigten hohe Übereinstimmungen, konnten aber auch zeigen, dass die Interrater Reliabilität abhängig von der Erfahrung der Durchführenden ist (Fox-Boyer und Meier, in Vorbereitung).

5. Schlussfolgerung

Die geforderten Kriterien zur Darstellung der Evidenz von Testverfahren für kindliche Aussprachestörungen sind zum jetzigen Zeitpunkt im deutschsprachigen Raum nicht ausreichend nachgewiesen. Lediglich für die PLAKSS-II konnten Objektivität, Inhaltsvalidität, diagnostische Validität (insbesondere da normiert) und Reliabilität nachgewiesen werden. Es ist zu überprüfen, ob die anderen existierenden Verfahren diese Kriterien erfüllen, da die dazugehörigen Handbücher bislang dazu nicht ausreichend Informationen bieten. Sollten die veröffentlichten Verfahren bereits gravierende Mängel im Bereich Inhaltsvalidität aufweisen, sollten diese überarbeitet oder nicht mehr verwendet werden. Da Kinder mit einer Aussprachestörung eine recht große Gruppe der Kinder mit Sprachauffälligkeiten im Kindergartenalter ausmachen, ist es sehr wichtig, dass Verfahren verwendet werden, die betroffene Kinder eindeutig als therapiebedürftig, und sprachgesunde Kinder eindeutig als sprachunauffällig identifizieren. Daher ist eine kritische Reflexion der vorhandenen Verfahren unabdingbar.

Literatur

Babbe T. (2003). Pyrmonter Aussprache Prüfung - Diagnostik von Ausspracheprüfungen bei Kindern (PAP). Köln: ProLog Therapie- und Lernmittel.

Bernhardt B.H. & Holdgrafer G. (2001). Beyond the basics

I: The need for strategic sampling for in-depth phonological analysis. *Language, Speech, and Hearing Services in Schools*, 32,18-27.

Beushausen U. (2008). Der Einsatz von standardisierten Tests in der Logopädie. *Forum Logopädie*, 22, 6-13.

Brunner E. (2015). Ausspracheprüfung für die Praxis (APP) - Österreich: Ein logopädisches Verfahren zur therapieorientierten Diagnostik kindlicher Aussprachestörungen. Wien: Neuroth Medical Division GmbH.

Crais E.R. (2011). Testing and beyond: Strategies and tools for evaluating and assessing infants and toddlers. *Language, Speech, and Hearing Services in Schools*, 42, 341-364.

de Langen-Müller U., Kauschke C., Kiese-Himmel C. et al. (2011). Diagnostik von Sprachentwicklungsstörungen (SES), unter Berücksichtigung umschriebener Sprachentwicklungsstörungen (USES) (Synonym: Spezifische Sprachentwicklungsstörungen [SSES]). Interdisziplinäre S2k-Leitlinie, Reg.-Nr. 049/006.

Dollaghan C.A. (2004). Evidence-based practice in communication disorders: What do we know, and when do we know it? *Journal of Communication Disorders*, 37, 391-400.

Dollaghan C.A. (2007). *The handbook of evidence-based practice in communication disorders*. Baltimore: Paul Brookes.

- Eisenberg S.L. & Hitchcock E.R. (2010). Using standardized tests to inventory consonant and vowel production: A comparison of 11 tests of articulation and phonology. *Language Speech and Hearing Services in Schools*, 41, 488-503.
- Felsenfeld, S., Broen, P.A. & MacGue, M. (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: Educational and occupational results. *Journal of Speech Language and Hearing Research*, 37, 1341-1353.
- Flipsen P, Jr., Ogiela DA (2015). Psychometric characteristics of single-word tests of children's speech sound production. *Language Speech and Hearing Services in Schools* 46:166-178.
- Fox-Boyer, A. & Meyer, S. (in Vorb.) Testgütekriterien für die PLAKSS-II.
- Fox-Boyer, A., Albrecht, K. & Clausen, M.C. (2018). Zur Problematik der mangelnden Kompatibilität psychometrischer und theoretischer Bedingungen diagnostischer Materialien. In Haring, R. & Siegmüller, J. (Hrsg) *Evidenzbasierte Praxis in den Gesundheitsberufen*, Untertitel: Chancen und Herausforderungen für Forschung und Anwendung. Heidelberg: Springer. S. 147-164.
- Fox-Boyer, A.V. (2014a). Aussprachestörungen. In: Grohnfeldt, M. (Hrsg) *Grundwissen der Sprachheilpädagogik und Sprachtherapie*. Stuttgart: Kohlhammer. S. 175-183.
- Fox-Boyer, A.V. (2014b). *PLAKSS-II Psycholinguistische Diagnostik kindlicher Aussprachestörungen II*. Frankfurt: Pearson Assessment.
- Fox, A.V. & Dodd, B. (2001). Phonologically disordered German-speaking children. *American Journal of Speech-Language Pathology*, 10, 291-307.
- Friberg, J.C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching and Therapy*, 26, 77-92.
- Hacker, D. & Wilgermein, H. (2002). *AVAK-Test. Analyseverfahren zu Aussprachestörungen bei Kindern*. München: Ernst Reinhardt.
- IEPMCS – International Expert Panel on Multilingual Children's Speech (2012). Multilingual children with speech sound disorders: Position paper. In: Research Institute for Professional Practice, Learning & Education (RIPPLE), Charles Sturt University, Bathurst, NSW, Australia.
- James, D.G.H. (2001). An item analysis of Australian English words for an articulation and phonological test for children aged 2 to 7 years. *Clinical Linguistics & Phonetics*, 15, 457-485.
- Kauschke, C. & Siegmüller, J. (2010). *Patholinguistische Diagnostik von Sprachentwicklungsstörungen*. München: Urban & Fischer / Elsevier GmbH.
- Kirk, C. & Vigeland, L. (2014). A psychometric review of norm-referenced tests used to assess phonological error patterns. *Language, Speech, and Hearing Services in Schools*, 45, 365-377.
- Kirk, C. & Vigeland, L. (2015). Content coverage of single-word tests used to assess common phonological error patterns. *Language, Speech, and Hearing Services in Schools*, 46, 14-29.
- Leitão, S. & Fletcher, J. (2004). Literacy outcomes for students with speech impairment: Long-term follow-up. *International Journal of Language and Communication Disorders*, 39, 245-256.
- Limbrick, N., McCormack, J. & McLeod, S. (2013). Designs and decisions: The creation of informal measures for assessing speech production in children. *International Journal of Speech-Language Pathology*, 15, 296-311.
- McCormack, J., McLeod, S., Harrison, L.J. et al. (2010). The impact of speech impairment in early childhood: Investigating parents' and speech-language pathologists' perspectives using the ICF-CY. *Journal of Communication Disorders*, 43, 378-396.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18, 5-11.
- Peña, E.D., Spaulding, T.J. & Plante, E. (2006). The composition of normative groups and diagnostic decision making: Shooting ourselves in the foot. *American Journal of Speech-Language Pathology*, 15, 247-254.
- Petermann, F. (2016). *Sprachstandserhebungstest für Kinder im Alter zwischen 3 und 5 Jahren (SET 3-5)*. Göttingen: Hogrefe.
- Schnitzler, C.D. (2015). *Schriftsprache und phonologische Verarbeitung bei Grundschulkindern mit im Vorschulalter überwundenen phonologischen Aussprachestörungen*.

rungen. Sprache Stimme Gehör, 39, 24-29.

Wagner, I. (2011). LOGO Ausspracheprüfung zur differenzierten Analyse von Dyslalien. Wildeshausen: Logo Verlag für Sprachtherapie.