

Probst, Holger

Schreibprobe Wissen & Denken (SDW)

Empirische Sonderpädagogik 4 (2012) 1, S. 22-46



Empfohlene Zitierung/ Suggested Citation:

Probst, Holger: Schreibprobe Wissen & Denken (SDW) - In: Empirische Sonderpädagogik 4 (2012) 1, S. 22-46 - URN: urn:nbn:de:0111-opus-92895

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Empirische Sonderpädagogik, 2012, Nr. 1, S. 22-46

Schreibprobe Wissen & Denken (SWD)

Holger Probst

Universität Gießen

Auf dem Weg klassischer Testkonstruktion entsteht ein linguistisch orientierter Test der basalen Rechtschreibkompetenz. Seine zwölf Untertests prüfen die Beherrschung/Nutzung aller relevanten linguistischen Wortsegmentierungen; sie unterscheiden sich in ihren Schwierigkeitsniveaus. Die Items und Subtests gehen aus Daten von 730 Grundschulern der Klassen 2-4 hervor; alle Gütekriterien sind erfüllt. Der Test bietet und empfiehlt kriteriale, sachbezogene Normen, die für jeden Subtest in Halbjahresintervallen das schwächste – „prekäre“ – Quartil ($PR < 25$) ausweisen. Dies geschieht nach Eingabe der 0/1-Lösungen in einem Excel-Template, das je Schüler individuelle Profile sowie ein Gesamtklassenprofil unter Vergleich mit der Mindestleistung für $PR > 25$ ausweist. An das individuelle Ergebnisprofil anschlussfähig sind mehrere vorliegende RS-Trainingsprogramme; ein spezifisch auf SWS abgestimmtes ist in Vorbereitung. Für den Gesamtwert liefert der Test auch PR-Normen nach sozialer Bezugsnorm.

Schlüsselwörter: Rechtschreibdiagnostik, Indikation von Rechtschreibförderung, kriteriale Diagnostik

The German writing test „Knowing and Thinking“

Based on the data of 730 elementary school students grade 2 to 4 a test for spelling/orthography is constructed. Item selection yields 12 scales on different difficulty levels, each of which represents a linguistically sound strategy towards correct writing, such as onset, rime, morph, long vs. short pronounced vowels. Concerning individual results the test offers criterial norms to define the lowest quartile (percentile < 25) in each subtest/ strategy, thus aiming at students at risk of orthographic failure. After entering 0/1 data into an excel-template the indication of remedial training results automatically for low performance students and subtests. Test indications comply with present training programs also based on linguistic word segments. In addition, the test offers percentile norms for the overall score.

Keywords: diagnostics of spelling/orthography, criterial diagnosis, diagnosis and intervention

In der Beherrschung der Schriftsprache kommt dem Lesen gegenüber dem (Recht-)Schreiben anerkanntermaßen die größere Beachtung und Bedeutung zu. So finden sich selten Einlassungen wie diese: „Die Fähigkeit, Wörter orthografisch korrekt zu schreiben, spielt eine bedeutende Rolle, wenn es ... um Schullaufbahntscheidung geht oder

die Versetzung... Aufgrund der zunehmenden Bedeutung der Schrift im privaten wie im beruflichen Alltag werden diese (Recht)schreib- H.P.) Kompetenzen immer wichtiger“ (Mayer 2009, S: 121). Damit kann – grundlegende – Rechtschreibkompetenz als Erkennungszeichen der Bildungsnähe und Voraussetzung für qualifizierte Ausbildungs-

gänge und schließlich als nach wie vor relevantes Lernziel gelten. Um den Lernweg zu basaler Orthographie zu ebnet empfiehlt sich auch hier der *Verbund von Test und Training*. In diesem Beitrag geht es um den Lenker eines solchen Tandems (mit „captain & stoker“).

Die Bezeichnung des Tests folgt der Auffassung, dass die Kompetenz zur Rechtschreibung (RS) sich zum einen Teil auf visuell eingeprägte und dann „gewusste“ Wörter stützt, zum anderen Teil aber auf Regelmäßigkeiten beruht, die der Schreiber in gelerntem Wortmaterial („taught words“) erkannt hat und nun – auch gegenüber neuen Wörtern – anwendet. Demnach beruht RS-Kompetenz auf *Rechtschreibwissen* wie auch auf *Rechtschreibdenken*, mit deren Hilfe sich auch unbekannte Wörter und Wort(teil)kombinationen korrekt schreiben lassen. Diese Sicht von RS-Kompetenz fußt auf linguistischen Unterteilungen des (deutschen) Wortmaterials, d.h. seiner Zerlegung in wiederkehrende Wortbausteine (Wortsegmente), deren Konstanz – mehr Regeln als Ausnahmen – das *RS-Denken* erst ermöglicht. Dass man mit einem Wortteil ein *Segment* vor sich hat, erweist sich, wenn dieses durch ein Element der gleichen Klasse ersetzt werden kann (Bußmann 1990, S. 670). So kann als Beispiel das Präfix ab-(schreiben) durch vor- (Element der gleichen Klasse) substituiert oder die phonotaktische Einheit zw- (erg) durch das Onset b- ersetzt werden. Nach dem Prinzip austauschbarer Laut- bzw. Schrifteinheiten sind es folgende Segmente, auf die sich RS-Kompetenz (und ihre Diagnostik) stützt: Grapheme (und ihre lautlichen Entsprechungen), Silben, Wortstämme und anhängende (gebundene) Morpheme sowie Anlautgruppen und Reime (Kluge, 1997), auch die Zerlegung von Komposita in Grund- und Bestimmungswort. Dies ist unser erster und wesentlicher Zugang zur RS-Kompetenz – Segmentierung.

Der zweite theoretische Zugang zur RS führt über die Phasen/Strategien der schriftsprachlichen Kompetenzentwicklung von

Kindern (Frith, 1985; Günther, 1986, 2004). Ihre anfängliche kurze Phase der Wortbild- (logographemische) Strategie lassen Kinder im Einschulungsalter mit dem Erkennen des Laut-Buchstaben-Bezugs hinter sich; sie schreiben nach der phonologischen Strategie und verfeinern diese bis zum richtigen Analysieren und Wiedergeben komplexer Lautpakete. In ihrer weiteren Entwicklung beachten Kinder orthografische Anforderungen wie Länge-/Kürzezeichen (z.B. h, ie, ck), die konstanten Wortstämme (mit ggf. Umlautung des Zentralvokals), und sie erkennen schließlich, dass gleich klingende Wörter je nach Sinn und Kontext unterschiedliche Schreibung verlangen. In diesem Verständnis von RS-Kompetenz konvergieren also entwicklungspsychologische und strukturalistische Theoreme.

Die Theorie der RS-Entwicklung ist bisher nicht so gezielt auf die linguistische Analyse zugegangen. Aber Schritte in die Richtung tat die Dual-Route-Theorie, die im aktuellen Vollzug der Verschriftung zwei Wege erkennt (Simon & Simon, 1973). Die erste Route vertraut der Phonem-Graphem-Zuordnung einer lautdarstellenden Graphie, also den in Klasse 1 gelernten Laut-Buchstaben-Zuordnungen. Auf der zweiten, „direkten“ Route reproduziert der Schreiber Graphemfolgen, also Wortbilder gemäß seiner Eintragungen im mentalen *orthografischen Lexikon*. Solche hat er oder sie, sobald Sehen und Schreiben über häufige Expositionen des Wortes zum Memorieren seiner Graphemfolge geführt haben. Orthografisch bedingte Abweichungen der Graphie vom lautbasiert zu Erwartenden hat der Schreiber mit besonderer Aufmerksamkeit markiert und differenziert so sein orthografisches Lexikon. In diesem Sinne erklärt Landerl (2000, S. 71): „Schwierigkeiten beim orthografischen Schreiben gründen sich auf eine Schwäche bei der Einspeicherung von Wortschreibungen im Gedächtnis.“ Die Zwei-Wege-Theorie (siehe auch Klicpera & Gasteiger-Klicpera, 1993, S. 77 ff.) sieht beim Schreiben also die sequentielle Fertig-

keit der Phonem-Graphem-Übersetzung oder das Abrufen orthografischer Gedächtniseinträge am Werk.

In der Lücke zwischen der Graphem-Synthese und den „eingespeicherten Wortschreibungen“ betont das SWD-Konzept den Zugriff auf orthografisch hilfreiche Wortsegmente, die recht verlässliche Umsetzungen von Wortklang in Wortschreibung nahelegen. Simon (1976) erweiterte die These vom memorierten Wortbild um des Schreibers Wissen um *Regelhaftigkeiten* der Schriftsprache, durch Heranziehen von *Information über Morpheme und Silben* des Wortes. Danach zieht der Schreiber Wissen um Wortteile dann heran, wenn er die Graphie weder kennt noch phonembasiert generieren kann. Regelmäßige Abweichungen (!) von der naiven Graphem-Phonem-Entsprechung wie z.B. bei der Auslautverhärtung werden zu Einträgen im orthografischen Lexikon und begründeten Schreibstrategien, die auf derartigen Segmenten beruhen (Greisbach, 2004; Kluge, 1997). Auf Rechtschreiben als Erkennen und Nutzen derartigen *Regelhaftigkeiten* (statt auf explizite, deklarative Rechtschreibregeln) setzt dann auch unsere Theorie der Rechtschreibkompetenz. In der Literatur mit Bezeichnungen wie „Sprachgefühl“, „Gespür für die korrekte Schreibung“ oder „Intuition“ mangelhaft angesprochen sind es implizite Regeln, die Kinder aus dem erfahrenen Schriftsprachmaterial ganz analog zum Regelsuchen beim Erwerb der mündlichen Sprache induktiv ableiten. Der Focus auf lautliche, orthografische und semantische Wortbausteine kann die weniger scharf definierten *Regelhaftigkeiten* der Schriftsprache präzisieren.

Weitere Stütze zieht die strukturalistische RS-Konzeption aus der Informationstheorie und aus der Informationsverarbeitung. Hier hilft das Konzept Redundanz weiter (die im Arbeitsspeicher aus chunks, Informations-„klumpen“ hervorgeht). Redundanz besteht dann, wenn ein Schriftwort mehr Zeichen (mehr Buchstaben = Graphen) enthält, als zu

seiner Lautierung oder Bedeutungserkennung nötig sind. Wortsegmente im orthografischen Lexikon bieten sich dem Schreiber an, wenn er zum Beispiel /fer-/ , /la:/ , /-tsion/ /- ə/ zu verschriften hat, nämlich diese: <ver-, -lah, -tion, -el>. Je inniger solche orthografischen Bausteine mit ihren Lautentsprechungen (konnektionistisch) verbunden sind, desto mehr Redundanz besteht, desto weniger (Phonem-Graphem-) Übersetzungsakte müssen die Schreibenden abarbeiten. Nur über die Bildung und Nutzung derartiger Superzeichen sind die Geschwindigkeit und die Leichtigkeit geübter Leser und Schreiber zu erklären.

In der Literatur finden sich nur wenige Hinweise auf diesen theoretischen Zugriff (Marx, 1997, S. 101; Grissemann & Roosen, 1989). Dabei ist dafürzuhalten, dass bereits mit dem Erwerb der Lautsprache innige Lautverbindungen, also redundante Hör- und Sprechkomplexe entstehen; phonotaktische Regeln definieren nicht-zulässige vs. „legale“ Phonemfolgen, für die später schriftsprachliche Entsprechungen, also Inhalte für das orthografische Lexikon zu erlernen sind. So erklären sich auch die weit reichenden Auswirkungen erstsprachlicher phonetisch-phonologischer Prägungen und auch metalinguistische Defizite auf den Schriftspracherwerb (Grimm, 1995).

Das Thema Rechtschreibung wird anders als Lesen in der psychologischen Literatur weniger beforscht (Schneider, 1997). Dabei besteht der Eindruck, dass Wortsegmente als Inhalte des orthografischen Lexikons wenig beachtet werden. Nur Morpheme werden als Gliederungshilfeprinzip genutzt (Kargl & Purgstaller, 2010); in der englischsprachigen Literatur fällt die Beachtung von „onset and rime“ (oder „rhyme“?) auf, weil sich im Englischen erst in der Einbettung im Reim eine verlässliche phonologische Konstanz der Vokalartikulation findet (Hines et al., 2007). Ganz anders werden in der pädagogischen Literatur und besonders in den dort erzeugten Trainingsmaterialien alle linguistisch defi-

nierten Sprachartikel ausdrücklich in Aufgabenformaten inszeniert (Heber & Cholewa, 2009). Wenn aber Wortverschriftung nach einem Bündel impliziter Regeln geschieht, so ist der Erwerb der Rechtschreibung auch ein Prozess des schlussfolgernden Denkens. Kinder produzieren bereits Ende der 1. Klasse Schreibweisen, die eindeutig über das Reproduzieren der Lernwörter hinausgehen, und lassen an Übergeneralisierungen erkennen, dass sie Hypothesen über lautabweichende Schreibungen aufgestellt haben. „Der selbständige Zugang von Kindern zur Schriftsprache und ihrer orthografischen Beherrschung ist also kognitiv bestimmt.“ (Röber-Siekmeyer, 1997, S. 44). Durch Vergleichen, also durch Informationsverarbeitung können Kinder erkennen, dass ein stimmhafter Konsonant als solcher erhalten bleibt, obwohl er im Auslaut „hart“ erklingt. Unter diesem Prinzip der morphologischen Konstanz (Stamm-schreibweise) können Kinder ebenfalls erkennen, dass ein Klang /oi/ von der Grundform /au/ abgeleitet und mit <äu> zu verschriften ist. Dieser logische Weg zur Rechtschreibung setzt allerdings ein höheres Maß an sprachlicher Bewusstheit voraus, unter der nicht mehr die Bedeutung eines Wortes, sondern seine etymologische Verwandtschaft Lernthema ist (Skowronek & Jansen, 2006, S. 129).

Nur zwei vorliegende RS-Tests beziehen sich m.W. erkennbar auf diesen theoretischen Hintergrund. Der Salzburger (Lese- und) Rechtschreibtest (SLRT; Landerl, Wimmer & Moser, 1997) unterscheidet über die Gesamtleistung hinausgehend zwei Fehlertypen, in denen Phasen der RS-Entwicklung aufscheinen. Es sind „schwerere“ Verstöße gegen eine lauttreue Verschriftung (Nicht-lauttreue Fehler) und schon „bessere“ Orthografische Fehler, die bei gelungener Lautdarstellung gegen die Orthographie verstoßen. Am weitesten in der Beherrschung der o.g. theoretischen Grundlagen fortgeschritten ist immer noch die Hamburger Schreibprobe (HSP; May, 1998). Eine gegensätzliche Blick-

richtung auf RS-Fehler verfiicht der Diagnostische Rechtschreibtest (DRT 2/3, Müller, 1996), der solche nach *funktionsätiologischen Kategorien* auf psych(olog)ische Mängel im Wahrnehmungs-, Memorier- oder Denksystem zurückführt (s. Valtin, 2009, S.50).

Zielsetzung/Fragestellung

Ziel der Konstruktion der SWD sind diese:

- Das Verfahren soll zeigen, welche Strategien des Zugriffs auf Segmentierungen und Nutzung impliziter Regeln ein Schreiber benutzt und beherrscht.
- Die Untertests/Skalen sollen die orthografische Tiefe deutscher Schriftsprache in zunehmend komplexeren Anforderungen repräsentieren.
- Der Geltungsbereich wird sich nach den Kompetenzen des Grundschulalters richten und soll für dessen Jahrgangsstufen Vergleichswerte liefern.
- Die Auswertung soll schnell und einfach sein und die Ergebnisse in leicht und direkt ablesbarer Form darstellen.
- Der Test soll alle Gütekriterien von Ökonomie bis Gültigkeit erfüllen.
- Tiefpunkte eines individuellen Profils sollen Förderschwerpunkte und -material indizieren.

Inhalt der Untertests

Die Auswahl der Items und ihre Zuordnung zu Subtests (= Skalen) orientierte sich an der oben dargestellten Theorie der Segmentierung deutscher Wörter. Die Itemformate konnten auf Erfahrungen mit einem Vorgänger – Inventar impliziter Rechtschreibregeln (liR, Probst 1989, 2001) – aufbauen und erfasssten RS-Phänomene, die in diversen Fördermaterialien trainiert werden. Vom ersten, leichtesten Subtest bis zu den schwierigsten loten die Subtests die orthografische Tiefe

deutscher Schriftsprache aus: Während bei Lauttreue mit „Tiefe Null“ direkte Phonem-Graphem-Entsprechung gilt, findet sich im mittleren Bereich die Konstanz von Wortteilen (wie Morphem, Länge-/Kürzezeichen) und noch tiefer unter der phonologischen Oberfläche die Bestimmung der angebrachten Schreibweise durch Semantik/Kontext. Die nun vorweg genommene Beschreibung und Begründung der SWD-Skalen folgt der oben dargelegten Theorie und der in Kap. 3 berichteten Skalenkonstruktion und Itemauswahl (siehe schon hier Tab. 1 im Anhang).

1. Vornamen. Der erste Pfad der Dual Route-Theorie, der indirekte, also synthetisierende Weg zum Verschriften basiert auf der einfachen Phonem-Graphem Übersetzung. Die Vornamen sind klangtreue Wörter (Synthesewörter, „Mitsprechwörter“), mit einfachen Konsonant-Vokal-Folgen (= K-V-K...). Für diesen leichtesten Test genügt die alphabetische (phonologische) Strategie. Übergeneralisierende orthografische Zutaten gelten nicht als Fehler. Nur bei schwachen Schülern finden sich nach Klasse 2 noch Lautverstöße oder unvollständige Lautdarstellungen.

Beispiele: Schreibe *Timo, Helena*. 5 Items

2. Reime und konsonantische Onsets sind deutsch-typische Hör- und Sprechmuster, tief eingravierte Phonemfolgen. Wörter/Wortstämme bestehen meist aus einem Konsonanten bzw. einer Konsonantengruppe im Anlaut (Onsets, siehe 3.) und der Reimgruppe, die mit dem Zentralvokal beginnt. Also: Wortstamm = Onset + Reim. Reime sind sinnfrei, aber klangkonstant. Die Reimwörter prüfen, ob das Kind die Schreibung einer Reimgruppe wahrnimmt und diese bei dem zu ergänzenden (Reim-)Wort nutzt und anwendet. Die Reimgruppen enthalten teils Kurz-, teils Langvokale, die zu erkennen sind. Kinder, die sich auf ihre lautorientierte Strategie „verlassen“, ignorieren die angebotene Information.

Beispiele: Vorgabe *Falle*, ergänze *H(alle)*. Vorgabe *Mahl* ergänze *k(ahl)*. 10 Items

3. Onsets sind Konsonanten/-häufungen/-verbindungen, hier sog. Konsonanten-Cluster am Wortanfang. Sie gelingen unter der lautlichen Strategie. Die Items sind leicht für Kinder, die solches aus ihrer mündlichen Sprache kennen, also Konsonantenfolgen analysieren und artikulieren können (Rigol, 1998, S. 20 ff.; Röber, 2011, S. 109). Sie sind schwerer für Kinder mit türkischer oder arabischer Erstsprache, die sich mit deutscher Phonotaxis (= Regeln zulässiger Lautfolgen), insbesondere den Konsonantenverbindungen schwer tun. Die Testwörter prüfen auch die Anlautungen *sp-* und *st-* und bieten die komplexesten Konsonantencluster; sie sind aber – für deutschsprachige Kinder – eher leicht.

Beispiele: Vorgabe *...erg*. Ergänze zu *Zwerg*. Vorgabe *...ücken*. Ergänze zu *pflücken*. 8 Items

4. Ableitungen. Das Erkennen und Beachten der Morphemgliederung deutscher Wörter führt zur korrekten Schreibung dieser sinntragenden Segmente, seien sie wie hier die Wortstämme oder die gebundenen Morpheme (s. 8. Affixe; Walter, 2001, S. 214 ff.; Mayer, 2009, S. 125). Die korrekte Schreibung des Zentralvokals (auch Zwiellaut) im Wortstamm bei Umlautung (*au-äu, a-ä*) gelingt bei Beachtung der klanglich eindeutigen Grundform. Analoges gilt für die Auslautverhärtung in Wörtern, die mit einem stimmhaften Konsonanten enden. Deren Schreibung gelingt beim Beachten der klanglich eindeutigen verlängerten (Plural-)Form. Diese Skala fordert mehr als die Lautanalyse, nämlich das Vergegenwärtigen der flektierten bzw. der Grundform des Wortes und ist bereits mittelschwer, liegt in mittlerer orthografischer Tiefe („Nachdenkwörter“).

Beispiele: Vorgabe *Kinder*, schreibe *Kind*. Vorgabe *laufen*, schreibe *Läufer*. 10 Items

5. Komposita sind Produkte der Worterzeugung durch Kombination (Reber, 2009, S. 77 f.). Sie sind mit der Phonem- und dann der Silbensegmentierung eine leichte Form der

Wortgliederung, weil ihre Zerlegung auch bei Kindern zu verständlichen semantischen Einheiten führt. Besonders die Nahtstelle, an der sich Bestimmungswort und Grundwort verbinden, ist eine RS-Klippe; bei nur lautlicher Orientierung verletzt der Schreiber die Intaktheit der Wortteile. Der Subtest ist mittelschwer, denn er verlangt hinter dem Lautstrom das Erkennen der Wortteile.

Beispiele: Schreibe nach Diktat *Wandtafel*; schreibe *Autounfall*. 7 Items

6. Lang-Vokal vs. Kurz-Vokal. In deutscher Sprache sind Lang- vs. Kurzvokale (auch bezeichnet als geschlossen vs. offen, gespannt vs. ungespannt) bedeutungsunterscheidend, sind also verschiedene Phoneme. Ihre Verschriftung unterscheidet dafür Länge- vs. Kürzezeichen (Röber-Siekmeier, 1998, S. 120 f.; Reber, 2009, S. 85-89). Hier ist zunächst zu bewerten, ob es dem Schreiber gelingt, den Klangunterschied in der korrekten Richtung zu verschriften (Schif – schief wäre gelungen, Schüff – schiff auch). Das Wahrnehmen des klanglichen Unterschiedes ist die Voraussetzung zur orthografisch korrekten Schreibung (zur Orthographie s. 11.). Besonders die germanisch-angelsächsischen Sprachen kennen die Bedeutungsopposition durch Lang- vs. Kurzvokal, darum ist die Unterscheidung besonders für süd-östlich geprägte Sprecher schwierig, deren Erstsprachen diese Opposition nicht kennen. Das lautliche Unterscheiden von Lang- und Kurzvokal ist eine zentrale Schwierigkeit jeglichen RS-Lernens und ist grundlegende Einsicht, ohne die Regelwissen der Verschriftung nicht genutzt werden kann (vertritt Brügelmann, 1994, S. 80).

Beispiel: Ergänze die Wörter: Das (*Schiff*) liegt (*schief*). 7 Items

7. Silben gelten als Verbindungsglied zwischen dem Prinzip der lautlichen und der morphologischen Verschriftung. Sie haben immer – nur – einen vokalischen Kern, der sprechmotorisch durch einen verstärkten Im-

puls der Atemmuskulatur betont wird (Blatt & Jarsinski, 2009; Walter, 2001, S. 174 ff.; Rigol, 1998). Das Erkennen von Silben wird sichtbar, wenn jeder Vokal (incl. Diphthonge) von dem anderen abgetrennt wird, gleichgültig, wo rechtschriftlich der Silbenschnitt liegt. Wer dies kann, lässt bei längeren Wörtern keine Silbe aus („Lokotive“).

Beispiele: Trenne in Silben: *Autobahnraststätte*, *Feuerwehrauto*. 5 Items

8. Affixe: Dies sind Präfixe und Suffixe am Wortstamm (= gebundene Morpheme; vulgo: „Vorsilben“ und „Nachsilben“) mit stets gleicher Schreibweise. Sie dienen der Wortbildung nach dem Prinzip der Affigierung. („Die Verwendung der lateinischen Begriffe ist wichtig, damit die Wortbildungsmorpheme eindeutig von der Sprech- und Schreibsilbe abgegrenzt werden.“! Blatt & Jarsinski, 2009, S. 108).

Beispiele: Ergänze die Wortteile (*un*)heim(*lich*). (*Ver*)käu(f*erin*). 12 Items

9. Homophone (Klang =/ Sinn ≠): Gleich klingende Wörter mit unterschiedlichen Schreibweisen haben dann auch unterschiedliche Bedeutungen. Hinter gleichem Klangbild muss der Schreiber gemäß Kontext einen semantischen Unterschied erkennen und entsprechend verschriften (Reber, 2009, S. 88). Der eher schwere Subtest verlangt die Kontext- und Sinnbeachtung. Allerdings erreichen die Schreiber hier leicht einige Treffer, wenn ein Paarling ein gelerntes Wort ist. In allen Items ist ein Paarling ein „Merkwort“, das einer selteneren Verschriftung folgt.

Beispiele: Ergänze in einem Lückentext *Meer – mehr*; *Häute – heute*. 8 Items.

10. Fremdwörter: Bei Wörtern lateinischen Ursprungs gelten nicht die deutschen Regeln des Laut-Buchstabenbezugs. Mit „Intuition“ und Analogiebildung kommen auch junge Schreiber hier zur korrekten Schreibung. Viele Fremdwörter sind bloße Merkwörter (oft mit hoher sozialer Valenz: Jeans,

T-Shirt, Inliner, Boots). Aber in manchen Fremdwortteilen erkennen schon Drittklässler Bekanntes und erkennen nun wiederum klangtreu geschriebene, unmarkierte lange /i:/ (in -ine, -iv, -il) sowie weitere lateinisch abstammende Partikel (Suffix -ion, sowie <v> für /w/. Röber-Siekmeyer traut und mutet schon Grundschulern das Fremdwörter-Schreiben zu (1997, S. 238 ff.).

Beispiele: Schreibe *Nation*, *Gardine*. 5 Items

11. Lang- vs. Kurzvokale orthografisch

korrekt: Wenn die auditive Unterscheidung von Lang- vs. Kurzvokalen gelingt (s.o. Punkt 6), ist darüber hinaus der richtige Einsatz der Länge- und Kürzezeichen gefragt, also die deutsche Konvention zur Markierung von Lang- vs. Kurzvokalen.

Beispiele: Die diktierten Lückenwörter (s. Punkt 6) orthografisch-korrekt schreiben. 14 Items.

12. Nachnamen sind Pseudowörter in deutsch-typischer Phonotaxis; sie korrekt (!, = normgemäß) zu schreiben gelingt nur beim Beherrzigen deutscher Konventionen zur Verschriftung. Dazu zählt auch das Schreiben klanglich uneindeutiger sog. Murmelsuffixe (wie -er, -el; das sind Endmorpheme mit dem schwa-Laut, dem tonlosen /-ə/ (Röber, 2011, S. 60-63, 148, 284; Mayer, 2009, S. 126).

Beispiele: Schreibe die Namen Kira Wocken, Helena Kriemel. 5 Items.

Methode

Die Konstruktion des Tests begann mit dem Schuljahr 2009 f. und der Erhebung einer Analysestichprobe aus Klassen 2-4. Zur Mitte desselben Schuljahres – im April/Mai 2010 – wurden weitere Klassenstufen 2-4 anderer Grundschulen untersucht. Eine dritte Erhebung geschah an einer weiteren Grundschule zu Beginn des Schuljahres 2010 mit Testwiederholung gegen Ende des Schuljahres, also Mai/ Juni 2011. Der Test geht also aus wiederholtem anwendungs- und beratungs-betonen Einsatz hervor, in dessen Zuge auf der Datenbasis von über 700 Kindern an Itemauswahl, an Gütekriterien, darunter auch Ökonomie der Anwendung gearbeitet wurde.

Für die Bezeichnung der Klassenstufen gilt folgende Bezifferung: Ende/Anfang meint üblicherweise die letzten/ersten drei Wochen des jeweiligen Schuljahres. 2.0 bezeichnet also Kinder nach ca. einem Jahr Unterricht. 3.5 bezeichnet Kinder in der Mitte der Klasse 3, also nach 2.5 Schuljahren. Klasse 4 ist nicht weiter differenziert.

Auswahl der Testaufgaben

Zu Beginn des Schuljahres 2009 f. wurden N = 283 Grundschüler der Klassen 2 bis 4 an vier Grundschulen mit der Analyseform der SWD getestet. Diese griff das Konzept des *Inventars impliziter Rechtschreibregeln* auf (s.o.) und erweiterte dessen Spektrum um die dort nicht erfassten linguistischen Einheiten bzw. Phänomene wie Anlautgruppen/Onsets, Komposita, Homophone und Fremd-

Tab. 1: Klassenstufe und deren Bezeichnung

| Ende 1/ Anfang 2 | Mitte Kl. 2 | Ende 2/ Anfang 3 | Mitte Kl. 3 | Ende Kl. 3 und Kl. 4 |
|---------------------|-------------|---------------------|-------------|-------------------------|
| 2.0 | 2.5 | 3.0 | 3.5 | 4 |

wörter. Die erste Testform hatte 137 Items in 14 Aufgabengruppen und einen Zeitbedarf von mitunter zwei Schulstunden (Spreyer, 2010).

Ziele der Itemanalyse und -selektion waren

- Subtests mit unterschiedlichen Schwierigkeitsniveaus entsprechend der orthografischen Tiefe der jeweiligen Anforderung/Strategie
- Items mit befriedigender Trennschärfe (r_{it})
- ein gewisse Streuung der Itemschwierigkeiten auch innerhalb jeder Skala
- Subtests/Skalen mit guter Konsistenz (Cronbach α) auch bei geringen Itemzahlen
- ein ökonomischer, in einer Schulstunde durchführbarer und leicht auswertbarer Test im o.g. Geltungsbereich

Nach der Analyse schieden aus: ein Untertest zur Groß-Kleinschreibung (die zum grammatischen Prinzip gehörend kaum sinnvolle Zusammenhänge zu den übrigen Subtests aufwies), ein Subtest zum visuellen Erkennen von Ganzwörtern sowie etliche Items mit unbefriedigenden Eigenschaften. Die Auswahl führte zu 98 Items, deren 78 für das 2. Schuljahr (Kl. 2.0 – 2.5) und deren 75 Items für 3.0 bis 4 vorgesehen sind.

Detaillierte Skalen- und Itemkennwerte sind im Anhang aufgeführt. Die Lösungshäufigkeiten S und deren Mittelwerte M_s zeigen, dass die linguistisch unterscheidbaren Anforderungen der Skalen zu unterschiedlichen Schwierigkeitsniveaus der Subtests führen. Dabei lassen sich mindestens befriedigende, meistens aber gute Skalenkonsistenzen (Cronbach α) erzielen. Im Falle der leichtesten Skalen klangtreue KV-Wörter (1. Vornamen) und 3. Onsets bleiben S hoch und r_{it} sowie α mäßig, weil unsere jüngsten Probanden (2.0), die Teil der Stichprobe sind, den lautdarstellenden Zugriff bereits gut beherrschen. Aber diese Aufgaben differenzieren schwache und jüngere Schreiber sehr wohl. Neben den Kennwerten wurden auch inhalt-

liche Aspekte des Wortmaterials beachtet; unter den Reimwörtern etwa sollten Minimalpaare mit vs. ohne Längen-/Kürzemarkierung vertreten sein.

Erneute Prüfung der SWD-Endform

Tabelle 2 zeigt die Ergebnisse der zweiten Itemanalyse aus 5/2010. Dazu dienten die Daten von $N = 253$ SchülerInnen der Klassen 2 und $N = 68$ der Klassen 3 und 4 um die Mitte dieser Schuljahre. Im Set der leichteren Subtests wie auch im Set der schwereren Subtests lösen die Schüler 56-58 % der jeweiligen gesamten Aufgaben. Der Prozentsatz gelöster Aufgaben ist für die beiden Klassenstufen ähnlich hoch, weil ab Klassen 3,0 schwerere Subtests zu bearbeiten sind (und leichtere entfallen). Die Auswahl insgesamt eher leichter Aufgaben führt zu einem Test, der die schwachen Schreiber feiner differenziert als die Spitzenkräfte. Dies entspricht der Zielsetzung, Förderindikation für Rechtschreibschwache zu leisten. In allen Klassenstufen decken die Skalen mit ihren mittleren Schwierigkeiten M_s ein Spektrum von rd. 90% bis rd. 20 % ab.

Tabelle 3 beleuchtet die Leistungsunterschiede zwischen Klassenstufe 2.5 und Klassenstufe 3.5. Während die Schüler sich in den lautanalyisierenden Zugriffen nur mäßig steigern – hier ist auch das Gros der Zweitklässler bereits weitgehend sicher –, haben die um ein Schuljahr weiteren Drittklässler bereits deutliche Fortschritte in der Beachtung von Stammprinzip, Wortsegmenten und Schreibkonventionen gemacht (bes. Lang- vs. Kurz-Vokal; s. Tab. 3). Jedoch stimmen die *Rangreihen* der Subtest-Schwierigkeiten von Klassenstufe 2.5 und 3.5 weitestgehend überein (Spearman $\rho = .94$). Änderungen der Schwierigkeitsfolge stellen sich hier und an weiteren Datensätzen nur zwischen ähnlich schweren Skalen ein.

Auf der neuen Datenbasis aus dem Schuljahr 2010 f. bestätigen sich die Item- und Ska-

Tab. 2: Kennwerte von Gesamt- und Subtests lt. erneuter Itemkontrolle 5/2010

| | 14 Klassen 2.5 N = 253 | | 4 Klassen 3- 4 N = 68 | | Anzahl Items |
|---------------------------|------------------------------|-------------------|-----------------------------|-------------------|-----------------|
| | Ms | Cronbach α | Ms | Cronbach α | |
| Vornamen | 0.92 | 0.25 | 0.87 | 0.44 | 5 |
| Reimwörter | 0.59 | 0.76 | 0.72 | 0.92 | 10 |
| Onsets | 0.76 | 0.55 | 0.53 | 0.94 | 8 |
| Ableitung | 0.69 | 0.82 | 0.77 | 0.61 | 10 |
| Komposita | 0.60 | 0.66 | 0.75 | 0.67 | 7 |
| Lang>kurz-Vokal | 0.36 | 0.76 | 0.54 | 0.81 | 7 |
| Silben trennen | 0.48 | 0.79 | 0.53 | 0.85 | 5 |
| Affixe | 0.52 | 0.74 | 0.73 | 0.63 | 12 |
| Homophone | - | - | 0.51 | 0.46 | 10 |
| Fremdwörter | - | - | 0.20 | 0.72 | 5 |
| Lang/kurz ortho.korr. | 0.22 | 0.78 | 0.48 | 0.84 | 14 |
| Nachnamen | - | - | 0.33 | 0.66 | 5 |
| Gesamt-Test Ms / α | 0.57 | 0.93 | 0.58 | 0.90 | 98 (78/75) |

| Subtest | Mittlere Leichtigkeit Ms | |
|--------------------|--------------------------|--------|
| | Kl.2.5 | Kl.3.5 |
| Vornamen | 0.87 | 0.93 |
| Reime | 0.76 | 0.86 |
| Ableitung | 0.68 | 0.85 |
| Onsets | 0.76 | 0.78 |
| Komposita | 0.59 | 0.78 |
| Affixe | 0.48 | 0.68 |
| Lang>kurz-Vokal | 0.35 | 0.54 |
| Silben | 0.48 | 0.53 |
| Klang=/Sinn \neq | - | 0.50 |
| Lang/kurz ortho | 0.22 | 0.48 |
| Nachnamen | - | 0.33 |
| Fremdwörter | - | 0.20 |
| Gesamt | 0.58 | 0.57 |
| N | 253 | 68 |

Tab. 3: Unterschiede zwischen Jahrgangstufe 2.5 auf 3.5, aber ähnliche Rangreihen der Subtest-schwierigkeiten

lenkennwerte (Tab. 2). Auch bei sehr kurzen Skalen mit nur fünf Items (Namen, Silben, Fremdwörter) liegen Cronbach α -Werte der Konsistenz für beide Klassenstufen im befriedigenden Bereich; die längeren Skalen zeigen sich sehr konsistent, so dass auch der Gesamttest gute interne Reliabilität ($r = .90/.93$) aufweist. Dieser Gesamtwert ist angesichts der hohen Zahl von Items nicht extrem hoch; dazu führen die unterschiedlichen Schwierigkeitsniveaus (S) der Subtests, weil sie die maximal mögliche Item-Interkorrelation deckeln. Erneut wird deutlich, dass sich entsprechend dem Testkonzept die Schwierigkeitsniveaus der Subtests deutlich unterscheiden. Sie rangieren von $S = .92$ für die Skala Vornamen bis in den Bereich um $S = .25$ für die Skala Pseudowörter/Nachnamen.

Gütekriterien

Relative Stabilität der Testwerte über drei Monate (Retest-Zuverlässigkeit)

Vier Grundschulklassen 3.5 liefern Daten von insgesamt $N = 58$ Schülern zur bloßen Testwiederholung nach einem Zeitraum von gut drei Monaten mit allfälligem Unterricht (von April bis Juli 2010; s. Tab. 4, Abb. 1). Die Retestzuverlässigkeit des Gesamttests beträgt $r_{tt} = .86$. Auch die z.T. sehr kurzen Skalen erreichen akzeptable Reliabilität.

Tab. 4.1: Retest-Zuverlässigkeit für Gesamt- und Subtests

| | N = 58 Kl. 3.5, r_{tt} |
|--------------------|-----------------------------|
| Vornamen | .55 |
| Nachnamen | .50 |
| Ableitung | .52 |
| Komposita | .48 |
| Lang>kurz | .62 |
| Lang/kurz orth. | .68 |
| Silben | .71 |
| Affixe | .54 |
| Klang=/Sinn \neq | .76 |
| Fremdwörter | .21 |
| SWD-Gesamt | .86 |

Im Rahmen einer Förderstudie (Hecht & Ennemoser, 2010) fallen im gleichen Zeitraum von drei Monaten Retestdaten für $N = 48$ Zweitklässler an, die für die Gesamt-SWD-Rohpunkte den Wert $r_{tt} = .88$ ergeben. Weitere Testwiederholungen von September 2010 bis Mai/Juni 2011 liefern die in Tabelle 4.2 dargestellten Konsistenz- (α) und Retestkoeffizienten (r_{tt}) des SWD-Gesamtwertes in fünf Lerngruppen/Klassen 2,3 und 4 einer Grundschule:

Tab. 4.2: Konsistenz- und Retestkoeffizienten der Testwiederholungen von September 2010 bis Mai/Juni 2011

| Klasse | N (für α) | α | N (für r_{tt}) | r_{tt} |
|--------|-------------------|----------|-------------------|----------|
| 2.5 | 59 | .94 | 17/6 | .82/.74 |
| 3.5 | 35 | .95 | 17/10 | .92/.89 |
| 4 | 36 | .96 | 17 | .89 |

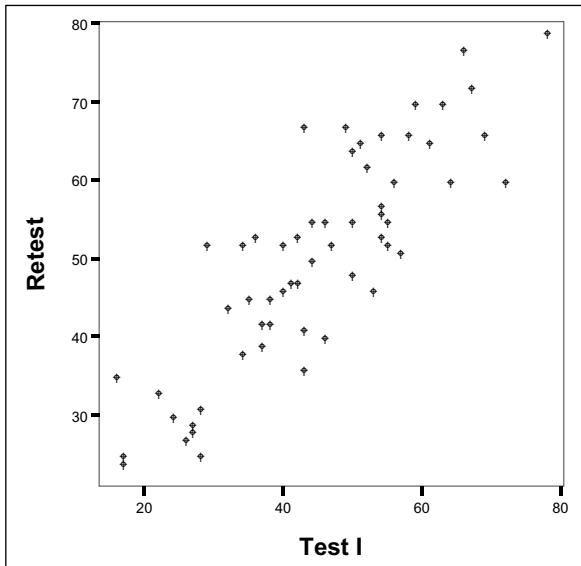


Abb. 1: Streudiagramm von $N=58$ Vor- und Retestwerten. Eine Gruppe extrem schwacher Schreiber verharrt bei sehr niedrigen Gesamt-Werten (SWD-Gesamtwerte, $r_{tt} = .86$)

Die mehrfach gefundene hohe Retest-Reliabilität r_{tt} über drei bis sechs Monate profitiert von der Heterogenität der Klassen und insbesondere dem Umstand, dass sehr schwache Schreiber konstant Schlusslichter bleiben und damit die Punktwolke in die Länge ziehen (siehe als Beispiel Abb. 1, Streudiagramm).

Die leichten Punktgewinne jener 58 Drittklässler im Retest sind signifikant für den Gesamt-Wert sowie für Affixe, Silben, Fremdwörter (t-Test für gepaarte Rohpunkt-Werte; Tab. 5). Der Fortschritt von 0,37 Standardabweichungen in drei Monaten entspricht der Faustregel, dass in einem Schuljahr die Schülerfortschritte eine Standardabweichung (1

Tab. 5: M , SD bei Testwiederholung, $N = 58$ Kl. 3. Rohwerte/richtige Schreibungen

| N = 58 | Test 1 | | ReTest (nach 3 Mo.) | |
|-----------------|--------|-------|---------------------|-------|
| | M | SD | M | SD |
| Vornamen | 4.39 | 0.92 | 4.36 | 0.93 |
| Nachnamen | 1.27 | 1.36 | 1.28 | 1.41 |
| Ableitung | 8.53 | 1.94 | 8.53 | 1.97 |
| Komposita | 5.19 | 2.14 | 5.38 | 1.78 |
| Lang>kurz | 3.23 | 2.27 | 3.81 | 2.10 |
| Lang/kurz orth. | 5.52 | 3.53 | 6.05 | 3.54 |
| Silben | 1.97 | 1.67 | 2.84 | 2.06 |
| Affixe | 7.56 | 2.70 | 8.83 | 2.56 |
| Klang/Sinn | 5.69 | 2.25 | 6.16 | 2.62 |
| Fremdwörter | 0.84 | 1.25 | 1.98 | 1.93 |
| SWD Gesamt | 44.19 | 14.21 | 49.38 | 13.94 |

SD) betragen, ihnen also eine Effektstärke von $ES = 1.0$ zukommt.

Kriteriale Gültigkeit

Stimmt der Gesamtwert der SWD-Ergebnisse mit den Zeugnisnoten überein, die Lehrer zum Schulhalbjahr ihren SchülerInnen erteilen? Die Noten (1-6) der Februar-Zeugnisse korrelieren mit im April 2010 erhobenen SWD-Ergebnissen. Zwei Klassen 2 und vier Klassen 3 einer Grundschule liefern Testwerte und Schulnoten.

(Diese Schule betreut Kinder von mehr als fünfzehn Nationalitäten in einem Quartier mit Sozialwohnungen, darin viele Familien in prekärer Lage. In zwei Klassen 3,5 wurden die von ihren Klassenlehrerinnen erfragten, kommentierten Noten korrigiert. In vier Fällen von „gut“ (2) war diese als „pädagogische Note“ mit dem Vermerk versehen „No-

te entspricht nicht der Benotung der Grundschule“ und wurde auf „3“ abgestuft. Die „Vier“ („4“) von vier anerkannten „LRS/LRR“-Schülern ging mit „5“ in die Berechnung, so auch die „4“ eines zur Sonderschulüberprüfung vorgesehenen Jungen. Plus- und Minus- (+, -) Vermerke wurde mit 0,3 Punkten numerisch umgesetzt.)

Daraufhin haben wir gültige Daten von $N = 42$ Zweitklässlern; für diese liegen eine Rechtschreib- und eine Lesenote vor. Für $N = 62$ Drittklässler haben wir nur eine nicht weiter differenzierte „Deutsch“-note; ihre Übereinstimmung mit SWD-Gesamtwert (% der erreichten Lösungen) zeigt Tab. 6. Die Korrelationskoeffizienten (Produkt-Moment) belegen hohe Validität der SWD am Kriterium Lehrerurteil.

Zum Bestimmen der externen Gültigkeit der SWD diente der Salzburger (Lese- und

Tab. 6: Koeffizienten der kriterialen Gültigkeit; SWD x Schulnoten/ SLRT

| | 2. Klassen N= 42 5/2010 | | 3. Kl. N = 62 5/2010 | 2. Kl. N = 18 9/2010 | 3. Kl. N =20 9/2010 | 4. Kl. N = 77 9/2010 | 3. Klassen N = 35 9/2009 |
|----------------------|-------------------------------|---------------|----------------------------|----------------------------|---------------------------|----------------------------|--------------------------------|
| | RS-Note | Lese- Note | „Deutsch“- Note | SLRT N+O Fehler | SLRT N+O Fehler | SLRT N+O Fehler | „Deutsch“- Note |
| Vornamen | -0.31 | -0.39 | -0.48 | -.17 | -.27 | -.47 | -.30 |
| Reimwörter | -0.68 | -0.64 | -0.56 | -.43 | . | . | -.59 |
| Onsets | -0.74 | -0.69 | -0.51 | -.54 | . | . | -.44 |
| Ableitung | -0.63 | -0.46 | -0.39 | -.46 | -.59 | -.72 | -.63 |
| Komposita | -0.58 | -0.60 | -0.63 | -.63 | -.38 | -.53 | -.63 |
| Lang>kurz | -0.65 | -0.62 | -0.65 | -.29 | -.67 | -.59 | -.66 |
| Silben trennen | -0.50 | -0.48 | -0.52 | -.64 | -.36 | -.49 | -.40 |
| Affixe | -0.62 | -0.53 | -0.54 | -.35 | -.75 | -.69 | -.63 |
| Klang≠Sinn | . | . | -0.68 | . | -.81 | -.73 | -.61 |
| Fremdwörter | . | . | -0.42 | . | -.62 | -.53 | -.38 |
| Lang/kurz ortho | -0.64 | -0.63 | -0.61 | -.45 | -.66 | -.71 | -.57 |
| Nachnamen | . | . | -0.48 | . | -.74 | N = 36 | -.35 |
| SWD-Gesamt | -0.86 | -0.80 | -0.78 | -.63 | -.88 | -.84 | -.73 |
| $\alpha < 0.05$ wenn | $r > .30$ | $r > .25$ | $r > .44$ | $r > .42$ | $r > .23$ | $r > .32$ | |

Rechtschreibtest (SLRT; Landerl, Wimmer & Moser, 1997), beide Tests zu Beginn des Schuljahres 2010 f. in den Klassen 2-4 zweier Grundschulen durchgeführt. Diese Gültigkeitskoeffizienten mit SWD-Gesamtwert (siehe Tab: 6) sind sehr (und nach psychologischen Maßstäben „unnötig“) hoch. Es erklärt sich vermutlich aus den hohen Streuungen der Schülerleistungen. Die Koeffizienten sind negativ, weil der SLRT Fehler zählt, deren Anzahl mit den Richtiglösungen der SWD umgekehrt korreliert; Gleiches gilt für die Schulnoten.

Tab. 6 differenziert das Bild und zeigt für die 12 Skalen, dass auch die Untertests signifikant mit der Schulnote korrelieren. An den Daten der Klassen 2 fällt auf, dass der Rechtschreibtest SWD fast ebenso gut mit der Leses-Note wie mit der RS-Note korreliert. Der Hauptgrund wird sein, dass hier bereits die beiden Noten mit $r = .82$ korrelieren. Lehrer differenzieren also weniger, als es Lese- vs. RS-Test tun. Aber auch bei Tests liegen die Lese-Rechtschreib-Korrelation um $r = .50 - .80$ (Klicpera & Gasteiger-Klicpera, 1993, S. 138) und beträgt auch in unseren Daten $r_{\text{SWD} \times \text{SLS}} = .62^{**}$ (Salzburger Lesescreening; Hecht & Ennemoser, 2010). Bereits in der Analyse-

stichprobe 2009 konnten die Deutschnoten von zwei 3. Klassen (zweier Grundschulen) erhoben werden; die Werte sind in Tab. 6, letzte Spalte angefügt.

Metrik – Kriteriale und soziale Normskalen

Aufbau des Tests: mit den Subtests steigende Anforderungen

Anders als bei Schulleistungstests üblich haben die Subtests der SWD unterschiedliche Schwierigkeitsniveaus. Natürlich kommt jede höhere Altersgruppe mit den steigenden Anforderungen besser zurecht. Abb. 2 stellt die mittleren Leichtigkeiten der 12 Skalen für drei Jahrgangsstufen dar. Die Niveaus der drei Linien zeigen, dass alle Subtests auf höherer Klassenstufe leichter fallen, dass aber die Steigerung der Anforderungen/Schwierigkeiten in den drei Jahrgangsstufen im Wesentlichen übereinstimmt (Daten aus 2011; drei Jahrgangsstufen einer Grundschule: Kl. 2.5; 3.5 bis 4; $N = 59/ 36/ 35$ resp.).

Abbildung 3 differenziert dieses Bild und zeigt die Verteilung der Itemschwierigkeiten je Untertest. Es basiert auf den gemeinsamen

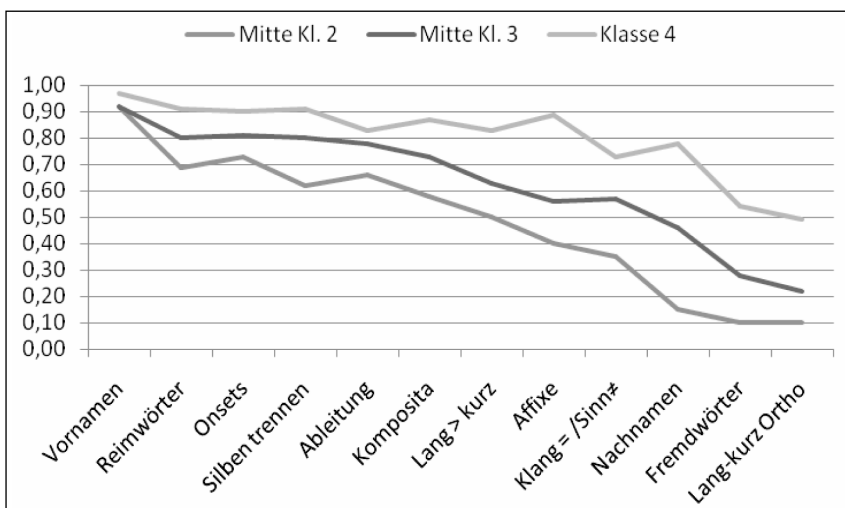


Abb. 2: Mittelwerte in % der Lösungen je Subtests für drei Klassenstufen (Daten 5/2011)

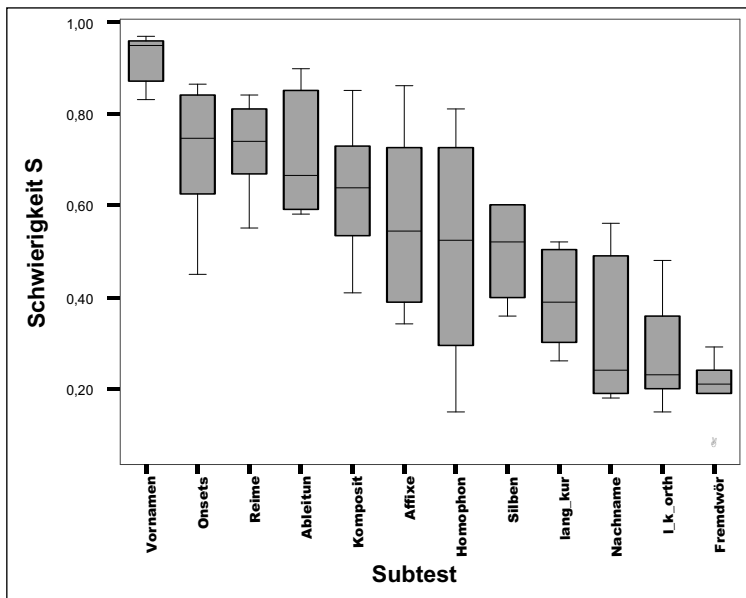


Abb. 3: Schwierigkeitsniveau und -breite der Subtests.

$N = 320$ Kl. 2.5 + 3.5 (Schuljahr 2010 f.)

S-Indizes für $N = 320$ Zweit- und Drittklässler. Die Mitte der Box markiert die mittlere Schwierigkeit (Median der S). Die Enden der Box fassen die mittleren 50% der Itemschwierigkeiten ein. Die Winker („whiskers“) umfassen 90% der S-Breite. Insgesamt macht Abb. 3 deutlich, dass die Untertests recht unterschiedlich schwer sind, dass aber auch jeder in sich eine Abstufung der Itemschwierigkeiten bietet. Dabei ermöglichen Subtests mit mittlerer Schwierigkeit die breiteste Variation ihrer Itemschwierigkeiten. Skalen mit extremer Schwierigkeit/Leichtigkeit tendieren zu links- bzw. rechtssteilen Verteilungen.

Kriteriale Bezugsnorm $PR < 25$

Dieses Verfahren möchte ausdrücklich diagnostischen Neulingen ein niederschwelliges, gleichwohl standardisiertes Angebot machen. Ausschließlich an sozialer Bezugsnorm orientierte Schulleistungstests mit oft testmethodisch überakelten Manualen erzeugen bei Grundschullehrkräften respektvolle Distanz, auch wenn dort Offenheit für den Sinn

der klassischen Gütekriterien besteht (Perleth, 1998). Die Übereinstimmung von Tests mit dem Lehrereindruck wird erwartet (z.B. Note, Ergebnisse aktueller Arbeiten). Bei erwartungswidrigen Ergebnissen forscht die Lehrkraft nach Erklärungen (z.B. der Schüler *hat abgeschrieben, hatte vor der Testung Stress, weckt durch mündliche Gewandtheit die Erwartung guter Deutsch-/RS-Leistung*). Eine bloße pauschale Bestätigung ihrer Schülereinschätzungen durch einen Standardwert im Testergebnis begeistert Lehrer nicht. Sie erwarten vielmehr weitergehende, differenzierende Information über das Gekonnte und die Defizite einer Schülerleistung und: Sie erwarten zu einem schlechten Testergebnis die Empfehlung einer Maßnahme, einer Intervention, die sie selbst oder die Helfer leisten können (Indikation durch Diagnose).

Diesen Wünschen will SWD entgegenkommen, zuerst durch eine direkte Metrik der Ergebnisdarstellung mittels kriterialer (sachbezogener) Normwerte. Kriteriale Normwerten ist leicht anzusehen, ob ein Schreiber eine getestete Strategie der Verschriftung beherrscht. Wenn er/sie z.B. fast

alle zehn gebotenen Affixe von ver- bis -ung korrekt schreibt und anfügt, kann man sicher sein, dass er/sie diese Partikel als solche erkennt und korrekt schreibt. Um dieses zu folgern ist es nicht nötig zu wissen, ob seine Mitschüler dieses können und wie er im Vergleich zu denen steht. Darum gibt der Test alle Subtest- und die Gesamtleistungen im Prozentsatz gekonnter Items an, genannt kriteriale Prozent/Sachprozent (kriteriale %). Das wirkt zunächst etwas primitiv, hat aber den Vorteil einer direkten Entsprechung im Grad der Beherrschung einer Strategie oder eines Stoffes. Lehrkräften ist diese Angabe geläufig, und nach dieser Logik beurteilen sie die Leistungen in Klassenarbeiten und informellen Lernkontrollen. Da es von SWD-Subtest 1 bis 12 zunehmend schwerer fällt und seltener wird, dass ein bestimmter Prozentsatz von Schreibungen richtig ausfällt, sinkt der Mindestwert hin zu den schwierigen Skalen; er steigt jedoch mit den Lehrjahren, also mit der

Klassenstufe. Die kriterialen Normwerte beachten diese beiden Faktoren. Tab. 7 gibt diese Mindestwerte als % und in () als Mindest-Itemzahl an. Ein Beispiel: Es ist für einen 3.0-Klässler nicht prekär, wenn er im Gesamttest nur 50% seiner Aufgaben schafft. Oder: Ein Schüler, der in Kl. 3.0 nur vier Affixe (= 33 %) richtig schreibt und anfügt, gehört zu den 25% der schwächsten Affixe-Schreiber.

Wie die PISA-Studie definiert SWD das erste, unterste Quartil mit PR < 25 als prekär. Die Mindestwerte für den PR >25 (vulgo: „grüner Bereich“) basieren auf jenen Daten, die im Zuge der Testanalyse und weiterer Anwendungen zu Schuljahrbeginn bzw. -mitte anfielen. Die Bestimmung der Mindestwerte ging aus von dem Sachprozentwert, dem in einer Klassenstufe der Sozial-Prozentrang 25 zukommt. Das ist ein Testwert (in kriterial %), den die 25% schwächsten Schreiber unterschreiten. Wegen der erheblichen Niveauunterschiede gleicher Jahrgangsstufen verschie-

Tab. 7: Mindestwerte für PR > 25 in % („grüner Bereich“) je Klassen-/Jahrgangstufe und Subtest (absolute Itemzahl)

| | Klassenstufe | | | | | Itemzahl |
|--------------------|--------------|--------|---------|---------|---------|----------|
| | 2.0 | 2.5 | 3.0 | 3.5 | 4 | |
| Vornamen | 80 (4) | 90 (4) | 100 (5) | 100 (5) | 100 (5) | 5 |
| Reime | 40 (4) | 60 (6) | 70 (7) | 80 (8) | 90 (9) | 10 |
| Onsets | 50 (4) | 60 (5) | 75 (6) | 90 (7) | 90 (7) | 8 |
| Ableitung | 30 (3) | 40 (4) | 50 (5) | 70 (7) | 80 (8) | 10 |
| Komposita | 30 (2) | 40 (3) | 60 (4) | 70 (5) | 85 (6) | 7 |
| Affixe | 25 (3) | 35 (4) | 40 (5) | 50 (6) | 65 (8) | 12 |
| Silben | 20 (1) | 40 (2) | 40 (2) | 60 (3) | 60 (3) | 5 |
| Lang>Kurz-Vokal | 0 | 15 (1) | 30 (2) | 40 (3) | 60 (4) | 7 |
| Klang/Sinn | - | - | 30 (3) | 40 (4) | 50 (5) | 10 |
| Lang/k. orthokorr. | 0 | 10 (1) | 15 (2) | 20 (3) | 30 (4) | 14 |
| Nachnamen | - | - | 0 | 0 | 20 (1) | 5 |
| Fremdwörter | - | - | 0 | 0 | 20 (1) | 5 |
| Mind. % gesamt | 30 | 43 | 40 | 52 | 63 | 98 |

dener Schulen (und sogar *einer* Schule) und wegen der großen Sprünge, wenn der Wert Sach-Prozent auf nur sieben oder fünf Items angewendet wird, war es notwendig, die Werte in Tab. 7 rational zu glätten und zu interpolieren. In dieser Form konturieren die Mindestwerte ein datenbasiertes Modell von hierarchischer Struktur und Zunahme von RS-Kompetenz (im Fall *Gesamtwert* ist zu beachten, dass ab Klasse 3,0 leichte Aufgaben entfallen und schwere dazukommen, so dass von 2.5 auf 3.0 die Leistungssteigerung nicht der Angabe der Mindestwerte in % sichtbar wird).

Die unterschiedlichen Anforderungen der Subtests rechtfertigen es, die jüngeren Klassen 2.0 und 2.5 mit den schweren Tests zu verschonen und ihnen nur die neun leichteren darzubieten. Wenn auch bis Kl. 3.5 noch keine Mindestleistungen in 10. Fremdwörter und 12. Nachnamen gefordert sind, erreichen hier etliche Kinder ab Klasse 3.0 sichere Lösungen, so dass es sinnvoll ist, von da an den Test komplett darzubieten. Man könnte nach zwei Schuljahren zur Zeitersparnis die drei leichten Tests auslassen; aber in allen Schulen fanden sich auch in Klassen 3.0 bis 4 förderbedürftige Kinder, deren Defizite bis in die Lauterfassung und Lautdarstellung zurückreichen. Insofern ist es für ein förderungsindizierendes Testverfahren probat, gerade geringen Kompetenzen auf den Grund zu gehen.

Soziale Bezugsnorm

Auch die soziale Bezugsnorm gibt wichtige Information. Sie macht zwar einen Umweg über M und SD der Vergleichsgruppe und verliert dabei die eigentliche Frage (etwa: Wie gut und sicher stellt jemand Lang- vs. Kurzvokale dar?) leicht aus den Augen. Aber alle Beteiligten – Lehrkraft, Schüler, Eltern – wollen auch wissen, wo ein Kind im Vergleich mit seiner Klasse(nstufe) oder Altersgruppe steht.

Die SWD bietet darum für den Gesamtwert auch soziale Bezugsnormen an und kommt dahin auf folgendem Weg: Ebenso wie für den sozialen PR 25 lassen sich für alle Perzentile die entsprechenden Testwerte (als Sachprozent) bestimmen. Sie finden sich in Tabelle 8, die 5er-Schritte für ausreichend differenziert hält. Die auf Sachprozent beruhenden Gesamt-Testwerte sind ebenso normalverteilt wie die Rohpunkte. Mithilfe einer Verteilung der Gesamt-% kann man sich klarmachen, dass ihr unterer/linker Ausläufer Kinder repräsentiert, die nur einen geringen Teil der Aufgaben korrekt schreiben.

Neben der Grenze, (nicht) zu den 25 % Schwächsten zu gehören, ist eine weitere PR-Grenze von praktischer Bedeutung. Es geht um die Bestimmung des unterdurchschnittlichen Bereichs. Er ist über die Standardabweichung definiert und geht aus dem PR hervor, welcher der Leistung von 1 Standardabweichung (-1 SD) unter dem Durchschnittswert entspricht. Dies ist der PR 16. Wer diesen PR mit seiner Testleistung nicht erreicht, fällt per Konvention in den unterdurchschnittlichen Bereich. Damit ist ein RS-Testwert unter PR 16 hinreichender Nachweis von Lernversagen und ein Kriterium für die Zuweisung von Förderbedarf Lernen. Um den Test auch in dieser Funktion nutzen zu können, hebt Tabelle 8 den PR < 16 hervor.

Es fallen Diagnose- und Beratungskontexte an, in denen auch unbedenkliche und überdurchschnittliche Leistungen nachgewiesen sein wollen. Dazu gibt Tabelle 8 für die Gesamttestwerte auch deren PR für die Gesamt-Skala bis PR 95 an und markiert den Median (= PR 50) als ein Maß der mittleren Leistung.

Auswertung; Eintrag in die Excel-Tabelle, Darstellung der Ergebnisse

Die Auswertung des Tests erstrebt ein differenziertes individuelles Ergebnis bei einfachs-

Tab. 8: Prozentrang (PR) in 5er-Intervallen für die kriterialen Testwerte „% geschafft“ je nach Klassen-/Jahrgangstufe

| Klassenstufe | | Kl. 2.0 | Kl. 2.5 | Kl. 3.0 | Kl. 3.5 | Kl. 4 |
|-----------------|----|----------------|----------------|----------------|----------------|--------------|
| N | | 114 | 272 | 104 | 117 | 129 |
| Mittelwert (AM) | | 0.44 | 0.56 | 0.55 | 0.60 | 0.70 |
| | | | | | | |
| Prozent-Rang | 05 | 0.17 | 0.25 | 0.20 | 0.32 | 0.40 |
| | 10 | 0.20 | 0.35 | 0.30 | 0.40 | 0.52 |
| PR < 16 | 15 | 0.23 | 0.40 | 0.34 | 0.44 | 0.57 |
| | 20 | 0.27 | 0.43 | 0.37 | 0.49 | 0.59 |
| Risiko PR < 25 | 25 | 0.30 | 0.45 | 0.42 | 0.52 | 0.62 |
| | 30 | 0.33 | 0.48 | 0.44 | 0.55 | 0.65 |
| | 35 | 0.38 | 0.50 | 0.49 | 0.56 | 0.66 |
| | 40 | 0.39 | 0.53 | 0.50 | 0.59 | 0.68 |
| | 45 | 0.43 | 0.55 | 0.53 | 0.61 | 0.70 |
| PR 50/Median | 50 | 0.46 | 0.57 | 0.55 | 0.62 | 0.72 |
| | 55 | 0.47 | 0.59 | 0.58 | 0.63 | 0.76 |
| | 60 | 0.49 | 0.60 | 0.62 | 0.64 | 0.78 |
| | 65 | 0.53 | 0.63 | 0.64 | 0.65 | 0.78 |
| | 70 | 0.55 | 0.65 | 0.66 | 0.68 | 0.80 |
| | 75 | 0.57 | 0.66 | 0.69 | 0.68 | 0.82 |
| | 80 | 0.59 | 0.69 | 0.71 | 0.71 | 0.83 |
| | 85 | 0.61 | 0.71 | 0.77 | 0.72 | 0.85 |
| | 90 | 0.63 | 0.75 | 0.80 | 0.75 | 0.88 |
| | 95 | 0.72 | 0.82 | 0.88 | 0.83 | 0.89 |

ter Auswertung, niederschwellig für alle Anwender. Eine programmierte Excel-Datei, ein Template (.xlt) leistet dieses (s. Abb.4). Eine Empfangsseite, die .xlt-Tabelle „Namen“ gibt Fenster für Namen, Junge/Mädchen, Klassen-/Vergleichsstufe. In Excel-Blatt 2 „Testeingabe“ trägt der Auswerter – links der Testbogen, rechts die Hand am Zahlenfeld der Tastatur – die richtige/falsche Schreibung mit 1/0 in einem Durchgang ein. Zur Orientierung steht über jeder Spalte das Item mit seiner laufenden Nummer. Alles Weitere ist programmiert: Auf den Button „Auswerten“ hin erscheinen in Blatt 3 namens „Profilliste“ au-

tomatisch die Kinder der Klasse/Testgruppe in absteigender Gesamtleistung. Farblich markiert erscheinen Testwerte (= %-Werte) mit PR < 25, immer mit Bezug auf die Mindestanforderung der gewählten Klassenstufe (s. Abb. 4).

Jede Zeile hinter dem Namen des Schülers gibt somit sein Profil über die leichten, mittleren, schwierigen RS-Anforderungen der Subtests von links nach rechts an. Bei schwachen Schreibern finden sich höhere %-Werte nur bei den leichten Skalen im linken Bereich einer Zeile. Gute Schreiber haben stets links hohe und auch in der Mitte und gar rechts

| | Vornamen | Reimwörter | Onsets | Ableitung | Komposita | Lang>Kurz-Voc | Silben trennen | Affixe | Klang = /Sinn# | Fremdwörter | Lang/kurz ortho. | Nachnamen | gesamt |
|---------|----------|------------|--------|-----------|-----------|---------------|----------------|--------|----------------|-------------|------------------|-----------|--------|
| Elias | 1.00 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.83 | 0.67 | 0.80 | 0.79 | 0.20 | 0.84 |
| Alexis | 1.00 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.58 | 0.50 | 0.60 | 0.86 | 0.40 | 0.81 |
| Franka | 1.00 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.83 | 0.50 | 0.00 | 1.00 | 0.60 | 0.81 |
| Julius | 1.00 | 0.67 | 0.63 | 0.90 | 1.00 | 0.86 | 0.80 | 0.50 | 0.75 | 0.20 | 0.79 | 0.60 | 0.72 |
| Maxim | 0.80 | 0.67 | 0.75 | 0.80 | 0.57 | 0.57 | 1.00 | 0.67 | 0.58 | 0.40 | 0.50 | 0.60 | 0.66 |
| Finn | 1.00 | 0.75 | 0.75 | 1.00 | 0.86 | 0.43 | 1.00 | 0.42 | 0.33 | 0.20 | 0.14 | 0.20 | 0.59 |
| Lotte | 1.00 | 0.75 | 0.63 | 0.60 | 0.57 | 0.57 | 0.80 | 0.92 | 0.33 | 0.20 | 0.29 | 0.20 | 0.57 |
| Tobias | 1.00 | 0.33 | 0.38 | 0.50 | 0.71 | 0.43 | 0.00 | 0.25 | 0.33 | 0.00 | 0.43 | 0.20 | 0.38 |
| Min 3.5 | 1.00 | 0.80 | 0.90 | 0.70 | 0.70 | 0.50 | 0.60 | 0.40 | 0.40 | 0.20 | 0.00 | 0.00 | 0.55 |

Abb. 4: Excel-Profil-Liste einer Lerngruppe 3.5. Erreichte kriteriale %. Prekäre Werte < Min für Kl. 3.5 werden grau unterlegt.

höhere/hohe Werte. Markierte, niedrige %-Werte deuten auf Förderschwerpunkte. Häufen sich markierte Werte spaltenweise, so ist dies ein Hinweis auf Rückstände der gesamten Klasse/Lerngruppe, oft RS-Schwerpunkte, die die Lehrkraft (noch) nicht behandelt hat. In der Klasse 3.5 des Beispiels (Abb. 4) wird die gehäufte Minderleistung in *Reime* und *Onsets* auf vernachlässigte auditive Differenzierung von Konsonantenverbindungen und Vokalen deuten und damit einen Förderschwerpunkt für die Lerngruppe anzeigen. Für Lotte und besonders Tobias zeigt das Profil eine breite/re Liste von Fehlerquellen, die hinter der schwachen Gesamtleistung sichtbar werden.

Eine weitere Excel-Prozedur stellt das Profil der Schüler als Säulendiagramm dar. Darin erscheint die persönliche Schülerleistung neben einer Säule, deren Höhe die Minimalanforderung angibt. Das Säulendiagramm emp-

fielt sich zusätzlich zur Tabelle für die Einzelfallberatung mit Eltern, Förderkräften oder mit dem Schüler selbst. Das Beispiel Hans in Abbildung 5 zeigt einen Schüler, der bei prekärer Gesamtleistung in den auditiven Zugriffen, der Lautdarstellung (Onset, Reim) und besonders beim Erkennen und Verschriften von Lang-/Kurz-Opposition weit hinter den Mindestanforderungen seiner 4. Klasse liegt.

Die auf den kriterialen Prozent basierten Werte bieten in der Excel-Auswertung auch die Möglichkeit, Klassenprofile und Prä-Post-Vergleiche grafisch zu veranschaulichen (Abb. 6, s. auch Abb. 7)

Die versprochenen Links des Testprofils zu entsprechenden Trainings lassen sich hier nur mit dem Verweis auf verfügbare RS-Fördermaterialien liefern. In ihrer linguistisch-segmentierenden Struktur sind diese Trainings konzeptionell den bisherigen RS-Tests

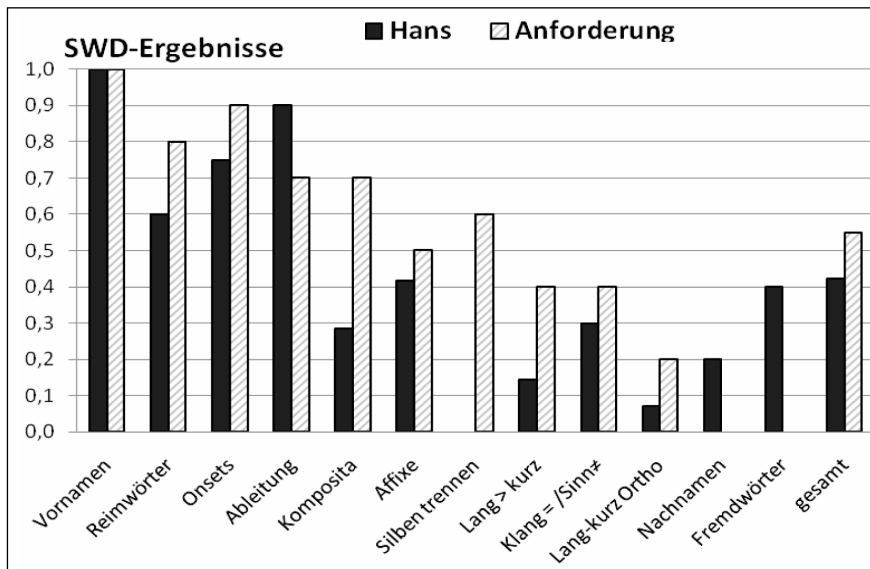


Abb. 5: Individuelles Ergebnis im Vergleich mit der Mindestanforderung für PR > 25

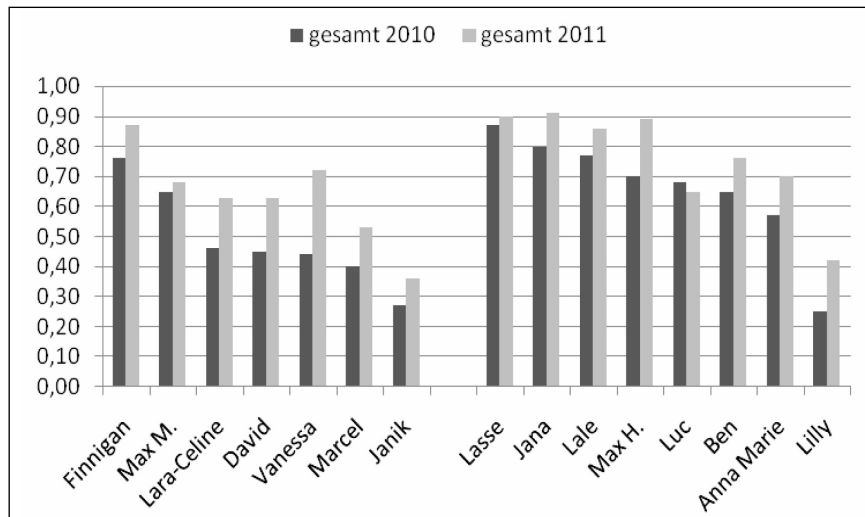


Abb. 6: Rückmeldung an Lehrkräfte: SWD-Gesamtleistung in einer jahrgangsübergreifenden Klasse 3 (links) und 4 (rechts) zu Beginn des Schuljahres 2010 und gegen Mitte/Ende Mai/Juni 2011.

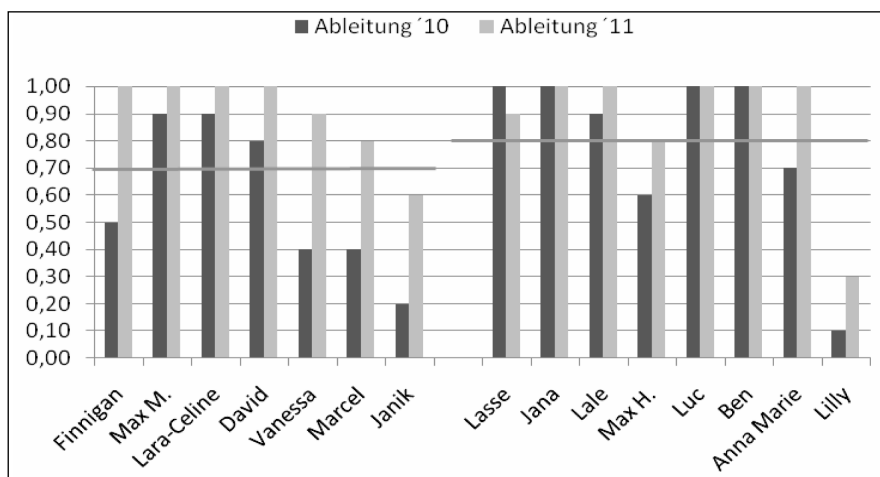


Abb. 7: Rückmeldung an Lehrkräfte: Verbesserung der Leistung im Subtest Ableitung in einer jahrgangsübergreifenden Klasse 3 (links) und 4 (rechts) zu Beginn des Schuljahres 2010 und gegen Mitte/Ende in Mai/Juni 2011. Über der Linie liegen Leistungen (% kriterial), die das Kriterium > PR 25 erreichen.

voraus. Das SWD-Profil liefert Hinweise auf die passenden Kapitel folgender RS-Trainings:

- Erst nachdenken, dann schreiben (Lautsch-Bregenzner, 2001)
- Lesen denken schreiben (Grissemann & Roosen, 1989)
- Rechtschreibleiter, bes. Band 3 (Hawellek, 2008)
- paLABra (PC-gestütztes Förderprogramm; Reber, 2006, 2009)
- Littera (Wildegger-Lack, 2003)
- Wortlistentraining (wlt, Balhorn, Harries & Schniebel, 1998),
- Morpheus (Kargl & Purgstaller, 2010)
- RS-Förderung durch Wissen & Denken (Hecht & Ennemoser, i. Vorb. 2012)

Diskussion

Die *Schreibprobe Wissen und Denken* versucht zwei Qualitäten von Schulleistungstests zu verbinden. Zum ersten ist es gelungen klassische Kriterien der Testtheorie zu erfüllen. So sind Standardisierung, Item-Schwierigkeiten und -Trennschärfen wie auch Skalen-

konsistenz zufriedenstellend; ebenso Ökonomie und Nützlichkeit. Die Gütekriterien Stabilität und externe Gültigkeit sind nachgewiesen. Die soziale Bezugsnormierung ist nur für die Region am Standort des Autors geleistet, aber damit als möglich demonstriert.

Ein zweiter Bereich angestrebter Qualitäten kommt mit dem Ziel, ein transparentes Verfahren für die Hand von Lehrkräften des Grundschulalters mit viel offensichtlicher Gültigkeit, mit niederschwelligem Zugriff und interventiver Relevanz anzubieten. Dies erzeugt an einigen Stellen Reibungen mit den Ansprüchen der klassischen Testtheorie:

Es ist nicht möglich (und theoretisch nicht denkbar) im gewählten Gültigkeitsbereich Untertests mit gleicher mittlerer Schwierigkeit zu konstruieren; die Verteilungen der extremen (leichten/schweren) Subtests können (anders als der Gesamtwert) nur schief sein. Damit sind die Voraussetzungen für einige Berechnungen nicht gegeben. Soweit es sich um Korrelationen handelt, fanden sich allerdings Werte gleicher Größenordnung für die parameterfreie Rangkorrelation.

Transparenz der Testergebnisse für Grund- (und Förder-)schullehrer kommt – nachdem diese die linguistisch motivierten Segmentierungen als Subtestinhalte akzeptiert haben – mit den direkten, kriterialen Werten zur Anforderungsbewältigung in Prozent (%) der richtigen Lösungen. Über die Skaleneigenschaften kriterialer Prozente ist wenig bekannt. Seit den 70er Jahren hat sich kriteriale (lernzielorientierte) Diagnostik (Klauer, 1987) nicht verbreiten oder gar durchsetzen können; aber Lehrkräften leuchten prozentuale Mindestangaben direkt ein – gerade wenn sie jahrgangsspezifisch moderiert sind. Dies zeigen die Erfahrungen mit den Lehrkräften im Zuge der hier berichteten Untersuchungen wie auch die tausendfachen (!) Erfahrungen mit einem ähnlich strukturierten Mathematiktest (Hinze & Probst, 2011). Unschön im Fall SWD sind die zwecks Testkürze geringen Itemzahlen einiger Skalen mit ihren Riesensprüngen auf der kriterialen Prozentskala (wenn 1 Item = 20%);

Die kriterialen Werte für die Mindestleistung sind offen für Korrektur und Justierung, sind jedoch wie hier berichtet hinreichend und konstitutiv für ein Modell (ähnlich Abb. 2), das Subtestanforderung/Leistung und Klassenstufe ins Verhältnis setzt. Die 25 % schwächsten Leistungen als prekär hervorzuheben folgt lediglich einer Konvention. B. Gasteiger-Klicpera (2010 mdl.) riet zur 10- oder gar 5-Prozentgrenze.

Zwar ist die Validität der SWD erwiesen; zwar erwiesen auch vorliegende RS-Trainingsprogramme ihre Wirksamkeit (Probst, 2009). Als valides Instrument für förderbedürftige Schüler und ihre RS-Problembereiche im Tandem von Test und Training müssen die SWD und ein ihr entsprechendes Training (Hecht & Ennemoser, in Vorb. 2012) sich noch beweisen.

Literatur

- Balhorn, H., Harries, B. & Schniebel, H. (1998). Wortlistentraining (WLT) 1-6. Hamburg: Verlag Pädagogische Medien.
- Blatt, I. & Jarsinski, S. (2009). Auswertung nach der Sprachsystematischen Rechtschreibdiagnose. In R. Valtin & B. Hofmann (Hrsg.), *Kompetenzmodelle der Orthographie* (S. 60-74). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Brügelmann, H. (1994). *Kinder auf dem Weg zur Schrift*. Konstanz: Faude.
- Bußmann, H. (1990). *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.
- Greisbach, M. (2004). Aufbau von Rechtschreibkenntnissen. In G. Lauth, M. Grünke & J. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 291-304). Göttingen: Hogrefe.
- Grimm, H. (1995). Spezifische Störungen der Sprachentwicklung. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 943-953). Weinheim: Psychologische Verlags Union.
- Grisseemann, H. & Roosen, H. (1989). *Lesen Denken Schreiben II – Förderung Lesen und Rechtschreiben*. Rangendingen: Lipura.
- Günther, K.-B. (2004). Kindliche Erwerbsstrategien beim Lesen- und Schreibenlernen. In A. Möckel, E. Breitenbach, W. Drave & H. Ebert (Hrsg.), *Lese-Schreibschwäche* (S. 27-53). Würzburg: Edition Bentheim.
- Günther, K.-B. (1986). Ein Stufenmodell der Entwicklung kindlicher Lese- und Schreibstrategien. In H. Brügelmann (Hrsg.), *ABC und Schriftsprache: Rätsel für Kinder, Lehrer und Forscher* (S. 32-54). Konstanz: Faude.
- Hawellek, T. (2008). *Rechtschreibleiter, Morphematischer Bereich, Stufen 12-16*. Oberursel: Finken-Verlag.
- Heber, S. & Cholewa, J. (2009). Sublexikalische Schreibdefizite bei deutschsprachigen Drittklässlern mit schwerer Entwicklungsdysgraphie: Welchen Effekt hat ein Training der

- Onset/Reim-Bewusstheit? Heilpädagogische Forschung. 25, S. 14-35.
- Hecht, T. & Ennemoser, M. (2010). Indizierte Rechtschreibförderung in Lehrer- vs. peergestützter Methode. *Unv. Ms., Justus-Liebig-Universität Gießen.*
- Hecht, T. & Ennemoser, M. (2012, i.V.). *Rechtschreibförderung durch Wissen und Denken.* Horneburg: Persen.
- Hines, S. J., Speece, D. L., Walker, C. Y. & DaDeppo, L. M. W. (2007). Assessing more than you teach: the difficult case of transfer. *Reading and Writing*, 20, 539-552.
- Kargl, R. & Purgstaller, C. (2010). *Morpheus - Morphemunterstütztes Grundwortschatz-Segmentierungstraining.* Göttingen: Hogrefe.
- Klauer, K. J. (1987). *Lehrbuch der Theorie und Praxis kriteriumsorientierten Messens.* Göttingen: Hogrefe.
- Klicpera, C. & Gasteiger-Klicpera, B. (1993). *Lesen und Rechtschreiben - Entwicklung und Schwierigkeiten.* Bern: Huber.
- Klicpera, C., Schabmann, A. & Klicpera-Gasteiger, B. (2007). *Legasthenie.* München: Reinhardt.
- Kluge, W. (1997). Die linguistischen Grundlagen der Orthografie und das Rechtschreiblernen in der Grundschule. In S. Feuchert & P. Büttner (Hrsg.), *Schreiben lernen - Schreiben lieben* (S. 93-120). Fernwald: Litblockin.
- Laetsch-Bregenzer, Ch. (2001). *Erst nachdenken, dann schreiben! Mit dem Rechtschreibtest „Inventar impliziter Rechtschreibregeln“ (liR) von H. Probst.* Donauwörth: Auer.
- Landerl, K. (2000). Der Salzburger Lese-Rechtschreibtest (SLRT). In M. Hasselhorn, W. Schneider & H. Marx (Hrsg.), *Diagnostik von Lese-Rechtschreibschwierigkeiten. Tests und Trends* (S. 63-79). Göttingen: Hogrefe.
- Landerl, K., Wimmer, H. & Moser, E. (1997). *Salzburger Lese- und Rechtschreibtest (SLRT).* Bern: Huber.
- Marx, H. (1997). Erwerb des Lesens und des Rechtschreibens: Literaturüberblick In F. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 85-111). Weinheim: Beltz.
- May, P. (1998). *Hamburger Schreib-Probe (HSP).* Hamburg: Verlag für pädagogische Medien.
- Mayer, A. (2009). *Gezielte Förderung bei Lese- und Rechtschreibstörungen.* München: Reinhardt.
- Müller, R. (1996). *Diagnostischer Rechtschreibtest 2-3.* Weinheim: Beltz.
- Probst, H. (1989). *Inventar impliziter Rechtschreibregeln.* Marburg: Philipps-Universität.
- Probst, H. (2001). *Inventar impliziter Rechtschreibregeln - liR. im Rechtschreibtraining „Erst nachdenken - dann schreiben!“ von Ch. Laetsch-Bregenzer.* Donauwörth: Auer
- Probst, H. (2009). *Präventive Rechtschreibförderung von Risikoschülern durch ihre Grundschullehrerinnen.* *Heilpädagogische Forschung*, 35, 202-212.
- Reber, K. (2006). *Metalinguistische Intervention mit dem Computerprogramm paLABra* In R. Bahr & C. Iven (Hrsg.), *Sprache Emotion Bewusstheit. Beiträge zur Sprachtherapie in Schule, Praxis, Klinik* (S. 380-388). Idstein: Schulz-Kirchner-Verlag.
- Reber, K. (2009). *Prävention von Lese- und Rechtschreibstörungen im Unterricht.* München: Reinhardt.
- Rigol, R. (1998). *Alphabet und Silbe.* In R. Weingarten & H. Günther (Hrsg.), *Schriftspracherwerb* (S. 19-35). Hohengehren: Schneider.
- Röber, C. (2011). *Die Leistungen der Kinder beim Lesen- und Schreibenlernen.* Hohengehren: Schneider.
- Röber-Siekmeyer, C. (1997). *Die Schriftsprache entdecken - Rechtschreiben im Offenen Unterricht.* Weinheim: Beltz.
- Röber-Siekmeyer, C. (1998). *DEN SCHBRISERIN NAS. Was lernen Kinder beim „Spontanschreiben“, was lernen sie nicht? Didaktische Überlegungen zum Verhältnis gesprochener und geschriebener Sprache, dargestellt an dem Problem der Wortabtrennungen.* In R. Weingarten & H. Günther (Hrsg.), *Schriftspracherwerb* (S. 116-150). Hohengehren: Schneider.
- Schneider, W. (1997). *Rechtschreiben und Rechtschreibschwierigkeiten.* In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie, Serie Pädagogische Psychologie* (Bd. 3, S. 327-363). Göttingen: Hogrefe.

- Simon, D. P. (1976). Spelling: a task analysis. In *Instructional Science*, 5, S. 277-302.
- Simon, D. P. & Simon, H. A. (1973). Alternative uses of phonemic information in spelling. *Review of Educational Research*, 43, S. 115-137.
- Skowronnek, H. & Jansen, H. (2006). Das Lesen und Schreiben erlernen. In A. Fritz, R. Klupsch-Sahmann & G. Ricken (Hrsg.), *Handbuch Kindheit und Schule* (S. 123-137). Weinheim: Beltz.
- Spreyer, L. (2009). „liR-R“ - Diagnostik basaler Rechtschreibregeln mit einer Revision des Inventar impliziter Rechtschreibregeln. Gießen: Staatl. Prüfungsamt.
- Valtin, R. (2009). Ein Wort schreibe ich mit dreierlei Orthographie. In R. Valtin (Hrsg.), *Kompetenzmodelle der Orthographie – empirische Befunde und förderdiagnostische Möglichkeiten* (S. 50-57). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Walter, J. (2001). *Förderung bei Lese- und Rechtschreibschwäche*. Göttingen: Hogrefe.
- Wildegger-Lack, E. (2003). *Littera – Metalinguistisches Schriftsprachtraining in sechs Stufen*. Germering: Verlag Wildegger.

Korrespondierender Autor

PROF. DR. HOLGER PROBST
 Justus Liebig-Universität
 Otto-Behagel-Str. 10 F
 35394 Gießen
 Holger.Probst@psychol.uni-giessen.de

Anhang

Item- und Skalenkennwerte aus Analyse I. 9/2009. Itemschwierigkeit S , deren Mittelwert M_s , Trennschärfen der Items r_{it} , Konsistenz der Skalen als Cronbachs α . $N = 261$

| 1 Vornamen | S | r_{it} |
|-------------------|----------|----------------------------|
| Kira | 0.96 | 0.38 |
| Timo | 0.96 | 0.31 |
| Heiko | 0.91 | 0.28 |
| Boris | 0.89 | 0.28 |
| Helena | 0.89 | 0.28 |
| M_s/α | .92 | .49 |

| 2 Reime | S | r_{it} |
|--------------------|----------|----------------------------|
| stehen - gehen | 0.83 | 0.78 |
| blutig - mutig | 0.76 | 0.54 |
| Platz - Satz | 0.63 | 0.82 |
| Bahn - Kahn | 0.63 | 0.78 |
| Halle - Falle | 0.62 | 0.78 |
| Mahl - kahl | 0.62 | 0.83 |
| Teller - Keller | 0.54 | 0.78 |
| Glocke - Locke | 0.52 | 0.69 |
| Tanne - Kanne | 0.51 | 0.79 |
| schmutzig - putzig | 0.49 | 0.64 |
| M_s/α | .62 | .92 |

| 3 Onsets | S | r_{it} |
|-----------------|----------|-----------------------|
| knobeln | 0.91 | 0.41 |
| Zwerg | 0.91 | 0.42 |
| schlürfen | 0.87 | 0.23 |
| Kreide | 0.87 | 0.42 |
| Sprache | 0.79 | 0.42 |
| prüfen | 0.79 | 0.34 |
| Strafe | 0.72 | 0.42 |
| pflücken | 0.49 | 0.34 |
| Ms/α | .79 | .67 |

| 4 Ableitung | S | r_{it} |
|--------------------|----------|-----------------------|
| Kind | 0.98 | 0.23 |
| Weib | 0.87 | 0.42 |
| Schlag | 0.81 | 0.32 |
| Grab | 0.79 | 0.27 |
| Fäuste | 0.76 | 0.49 |
| Dächer | 0.75 | 0.39 |
| Läufer | 0.68 | 0.64 |
| Bäumchen | 0.66 | 0.45 |
| Räuber | 0.60 | 0.60 |
| Käufer | 0.51 | 0.48 |
| Ms/α | .74 | .77 |

| 5 Komposita | S | r_{it} |
|--------------------|----------|-----------------------|
| Wandtafel | 0.79 | 0.47 |
| Gemüse | 0.72 | 0.47 |
| Hüpfburg | 0.71 | 0.54 |
| Holzzaun | 0.70 | 0.54 |
| Blitzschlag | 0.66 | 0.47 |
| Kunststück | 0.48 | 0.59 |
| Autounfall | 0.48 | 0.56 |
| Ms/α | .65 | .79 |

| 6 Lang > Kurz-Vokal | S | r_{it} |
|-------------------------------|----------|-----------------------|
| Müller-Mühle | 0.42 | 0.50 |
| fehlt-Fell | 0.42 | 0.50 |
| Schiff-schief | 0.41 | 0.49 |
| fühlt-füllt | 0.28 | 0.45 |
| Lamm-lahm | 0.27 | 0.45 |
| Höhle-Hölle | 0.22 | 0.42 |
| kullern-Kuhle | 0.19 | 0.39 |
| Ms/α | .32 | .85 |

| 7 Silben | S | r_{it} |
|---------------------|----------|-----------------------|
| Schokoladentafel | .68 | 0.68 |
| Autobahnraststätte | .68 | 0.67 |
| Abenteuerspielplatz | .64 | 0.67 |
| Ruderbootverleih | .63 | 0.68 |
| Feuerwehrauto | .62 | 0.74 |
| Ms/α | .65 | .87 |

| 8 Affixe | S | r _{it} |
|------------|------|-----------------|
| unheim | 0.75 | 0.36 |
| heimlich | 0.94 | 0.29 |
| Vorschrift | 0.75 | 0.53 |
| schriften | 0.60 | 0.15 |
| Verkäuf | 0.78 | 0.53 |
| käuferin | 0.63 | 0.00 |
| veränd | 0.76 | 0.36 |
| ändert | 0.74 | 0.15 |
| Verabred | 0.76 | 0.45 |
| abredung | 0.69 | 0.36 |
| auffäll | 0.72 | 0.06 |
| fällig | 0.66 | 0.27 |
| Ms/α | .73 | .63 |

| 9 Homophone | S | r _{it} |
|---------------|------|-----------------|
| Meer - mehr | 0.49 | 0.52 |
| Boot - bot | 0.36 | 0.45 |
| Häute - heute | 0.34 | 0.48 |
| Rad - Rat | 0.27 | 0.42 |
| hohl - holen | 0.23 | 0.36 |
| fiel - viel | 0.18 | 0.41 |
| Ms/α | .31 | .71 |

| 10 Fremdwörter | S | r _{it} |
|----------------|------|-----------------|
| Nation | 0.33 | 0.49 |
| Gardine | 0.29 | 0.46 |
| Motiv | 0.29 | 0.28 |
| Ventil | 0.28 | 0.63 |
| Reaktion | 0.25 | 0.61 |
| Ms/α | .29 | .73 |

| 11 Lang-kurz orthogr.korr. | S | r _{it} |
|-------------------------------|------|-----------------|
| Schiff | 0.53 | 0.56 |
| schief | 0.54 | 0.50 |
| Lamm | 0.68 | 0.18 |
| lahm | 0.37 | 0.51 |
| fehlt | 0.51 | 0.55 |
| Fell | 0.71 | 0.39 |
| Müller | 0.78 | 0.31 |
| Mühle | 0.43 | 0.63 |
| Höhle | 0.31 | 0.49 |
| Hölle | 0.53 | 0.37 |
| fühlte | 0.44 | 0.61 |
| füllte | 0.29 | 0.52 |
| kullern | 0.41 | 0.56 |
| Kuhle | 0.24 | 0.52 |
| Ms/α | .48 | .84 |

| 12 Nachnamen | S | r _{it} |
|--------------|------|-----------------|
| Rahn | 0.49 | 0.49 |
| Wocken | 0.49 | 0.29 |
| Holling | 0.18 | 0.57 |
| Schwenner | 0.24 | 0.45 |
| Kriemel | 0.19 | 0.30 |
| Ms/α | .33 | .66 |