

## Wie wirken verschiedene Methoden des Rechtschreibunterrichts?

In dieser überarbeiteten Fassung eines früheren Beitrags (Brügelmann 2013) werden Befunde aus empirischen Untersuchungen zu verschiedenen didaktischen Ansätzen zusammengefasst. Dabei stehen Studien aus Deutschland im Vordergrund, da sich Sprache, Unterrichtstradition und Schulsystem als bedeutsame Randbedingungen erwiesen haben.

Wie sich zeigen wird, ist das Resultat für diejenigen wenig befriedigend, die sich einfache Antworten erhoffen. In der Didaktik gibt es keine Entscheidungsexperimente, deren Befunde erlauben, Methoden als generell »wirksamer« zu beurteilen. Je nach Konzept, in das sie eingebettet werden, und je nach Kontext, in dem sie umgesetzt werden, streuen die Effekte jeweils breit (vgl. grundsätzlich zu diesem Problem: Brügelmann 2015a, Kap. 2). Obwohl in der Regel standardisiert erhoben sind die im Folgenden berichteten Daten deshalb hoch interpretationsbedürftig. Unter diesem Vorbehalt können sie aber auch für die Unterrichtspraxis anregend sein – und vor allem einer naiven Methodengläubigkeit entgegenwirken.

### Wortschatzübung

Dieser Ansatz ist am umfassendsten untersucht. Schon früh haben Studien gezeigt, dass die Übung von einzelnen Wörtern deren Richtigschreibung fördert. Das systematische Training einer Wörterliste in 3. Klassen über ein halbes Jahr hinweg reduzierte die Fehlerquote bei diesen Wörtern um gut die Hälfte (Plickat / Lüder 1979, 128 – 130; Beck / Eisenhauer 1979, 139 – 146). Auch bei ungeübten »Transferwörtern« (mit gleichen Rechtschreibbesonderheiten) nahmen die Falschreibungen deutlich ab. Zehn Monate später allerdings zeigten sich Vorteile nur noch zugunsten der Übungswörter selbst. Wiczerkowski u. a. (1979a + b) fanden in ihrer Untersuchung des »Wortlisten-Trainings« von Balhorn ebenfalls nur einen geringen Transfer-effekt auf ungeübte Wörter.

Dabei macht es nach Nickel (1979, 146 – 164) keinen Unterschied, ob die Wörter einzeln oder in Texten geübt werden. Bedeutsam ist aber der persönliche Bezug zu den Wörtern bzw. dem Erfahrungsfeld, zu dem sie gehören, sei es wegen der motivationalen Ladung, sei es wegen des Übungseffekts durch besonders häufige Nutzung. Dies konnten May (1994) und Richter (1996) beim Vergleich von Jungen- und Mädchenwörtern nachweisen. Insofern spricht viel dafür, den Übungswortschatz von den Kindern individu-

ell bestimmen zu lassen (vgl. den »interessengeleiteten« Ansatz von Richter 1998). Dies auch deshalb, weil sich nach dem Prinzip der Häufigkeit nur ein kleiner Wortschatz als »gemeinsam wichtig« auszeichnen lässt (vgl. Brügelmann 2015b).

Die Hoffnung, dass es die schulische Übung einzelner Wörter ist, die deren Fehlerquote senkt, ist ebenfalls einzuschränken. Zwar fanden Brügelmann u. a. (1994d, 175) bei der Auswertung des Schreibvergleichs BRDDR, dass in freien Texten die Fehlerquote von Wörtern aus dem in der DDR intensiv geübten Mindestwortschatz mit 3 – 4 % deutlich geringer war als bei Wörtern außerhalb des Mindestwortschatzes mit 18 – 19 %. Aber fast dieselbe Differenz fand sich mit 5 – 6 % zu 22 – 23 % auch bei den Kindern aus den westlichen Bundesländern, obwohl das Üben eines festen Wortschatzes dort einen geringeren Stellenwert hatte und die Zusammensetzung der Grundwortschätze von Bundesland zu Bundesland variierte. Noch bedeutsamer: Bereits innerhalb des Mindestwortschatzes sprang die Fehlerquote oberhalb des 430. Wortes (in der Häufigkeitsfolge) von ~5 % auf ~20 %. Dieser Sprung ließ sich übrigens nicht nur im freien Text, sondern auch im Diktat beobachten.

Gegen die Hoffnung mancher PolitikerInnen, dass eine *landesweit verbindliche* Wörterliste den Rechtschreibunterricht und über ihn die Rechtschreibentwicklung steuern könne, sprechen auch die länderbezogenen Analysen der IGLU-Studie 2001 (vgl. Valtin u. a. 2004, 158ff., zum Teil eigene Berechnungen): Zwar schreiben Thüringer Schüler/innen von den Wörtern aus der landesweiten Empfehlungsliste 62 % richtig – also mehr als bei den Wörtern außerhalb (49 %). Aber dies ist auch in anderen Bundesländern der Fall, die keine Wörterliste vorgeben. Bayerische Schüler/innen schreiben innerhalb der dort zur Übung vorgeschriebenen Wortauswahl ebenfalls nur 59 % der Wörter richtig – außerhalb des Grundwortschatzes aber sogar mehr, nämlich 68 %. Auch diese Differenzen finden sich ähnlich in den anderen Bundesländern. Fazit: es scheint eher die generelle Häufigkeit (des Schreibens) von Wörtern – und vermutlich auch deren orthographische Schwierigkeit – als ihre Übung im Unterricht zu sein, die zur Rechtschreibsicherheit beiträgt.

Daneben ist es die persönliche Wichtigkeit bei der Verwendung von Wörtern, die die Richtigschreibung stützt (s. Richter / Brügelmann 1994, 121 ff.; May 1994a). Dies spricht ebenfalls gegen die Verordnung eines für alle gleichen Übungswortschatzes.

Zu bedenken auch ein weiterer Befund aus dem »Schreibvergleich BRDDR«: Zwar waren die Viertklässler/innen der ehemaligen DDR – in einem Diktat von Wörtern aus dem dort systematisch geübten Mindestwortschatz – mit durchschnittlich nur 12 falschen Wörtern auf 100 den Kindern aus Westdeutschland (16 Fehler) und aus Schweizer »Lesen durch Sch-

reiben-Klassen« (18 Fehler) deutlich überlegen. In einem freien Text waren die Unterschiede aber nur noch statistisch, nicht mehr inhaltlich signifikant: 7 Fehler gegenüber 9 in den beiden anderen Gruppen (vgl. Brügelmann u. a. 1994b, 137, zu dieser als »Laufstalleffekt« des Übens bezeichneten fehlenden Übertragung vom Diktat auf freie Texte). In spezifischeren Auswertungen zeigte sich, dass sich auch das Fehlerprofil der ehemaligen DDR-Klassen (trotz des ganz anderen Rechtschreibunterrichts) nicht von dem der beiden anderen Gruppen unterschied (Brügelmann u. a. 1994d, 174 – ähnlich für eine Auswertung nach Rechtschreibstrategien May 1991). Im Übrigen widersprechen die Ergebnisse der Analphabetismus-Studie leo auch grundsätzlich der Annahme, dass der Schriftsprachunterricht in der DDR – gemessen an den nachschulischen Leistungen – erfolgreicher war (Grotlischen u. a. 2012, 36 f.).

Umstritten ist, **ob häufiges Lesen die Rechtschreibsicherheit fördert** (vgl. schon Bryant/Bradley 1980). Gegen einen solchen Transfer spricht, dass die Aufmerksamkeit beim Lesen auf den Inhalt der Texte und nicht auf die Form der Wörter gerichtet ist. Andererseits wird beim Lesen die implizite Assoziation häufiger Buchstabenfolgen im inneren Lexikon gefestigt. Diese Automatisierung könnte sich auch positiv auf ihren Abruf beim Schreiben auswirken (vgl. das Modell entfalteter Rechtschreibkompetenz bei Scheerer-Neumann i. d. B.). Eine Untersuchung von Siekmann (2011; 2013) fand keinen wortspezifischen Transfer vom Lesen auf das Rechtschreiben. Auch in früheren Studien hat sich die bewusste Auseinandersetzung mit der Schreibweise des einzelnen Wortes gegenüber bloßem Abschreiben oder mehrfachem Lesen als effektiver erwiesen (vgl. Peters 1974 [zit. nach Downing 1979, 75f.]; van Doorn-van Eijnsden 1984; Bosman/de Groot 1992, 280f., 287; zu ähnlichen Befunden aus neueren Studien wie DESI vgl. Siekmann 2013). Beck u. a. (2009) haben zudem gezeigt, dass sich die Gruppen der schwachen Leser/innen und Rechtschreiber/innen nur begrenzt decken. Zusammenfassende Etiketten wie »LRS« oder »Legastheniker« sind insofern für die Förderung nicht hilfreich. Vielmehr muss individuell erhoben werden, wo genau ein Kind besondere Probleme hat: beispielsweise eher beim Verschriften unbekannter Wörter oder beim Abruf häufig verwendeter Wörter aus dem Gedächtnis, bei der Beachtung der Wortverwandtschaft für die Ableitung von Schreibungen oder bei der Nutzung sprachstatistischer Faustregeln, beim Nachschlagen im Wörterbuch oder beim selbstständigen Üben besonders fehlerträchtiger Wörter.

## Phänomen- bzw. regelorientierter Unterricht

Eine Reihe von Studien konzentriert sich auf die Aneignung spezifischer Phänomene der Rechtschreibung, für die bestimmte Abfolgen des Erwerbs postuliert werden (zusammengefasst nach Risel 2008, 68–70): zur Konsonantenverdopplung nach Kurzvokal (Günther 2006), zur Verschriftung des /f/-Lauts (Thomé 1999), zur Verwendung von Satzzeichen (Eichler/Küttel 1993; Afflerbach 1997), zur Getrenntschreibung (Bredel 2006) und zur Trennung bei Zeilensprung (Geilfuß-Wolfgang 2004). Zur Groborientierung sind die Befunde durchaus hilfreich. Aber auch diese Erwerbsstufen dürfen nicht zu rigide verstanden werden. Insofern sind Altersangaben problematisch, die die große Streuung um die Durchschnittswerte verdecken. Und auch bei den Durchschnittswerten muss offen bleiben, inwieweit sie durch die Eigendynamik des kindlichen Rechtschreiberwerbs oder doch eher durch Unterrichtstraditionen bestimmt sind. Sie können also nicht unbesehen als Richtlinie für den Aufbau des Rechtschreibunterrichts genommen werden.

Phänomen- bzw. regelorientierte Ansätze konzentrieren Rechtschreibübungen auf einzelne Rechtschreibbesonderheiten (z. B. das <h> nach Langvokal) bzw. die Vermittlung von explizit formulierten Regeln (z. B. die Stammkonstanz bei der Deklination und Konjugation).

So übte Dumke (1979, 84–93) – gestreckt über acht Monate – in zweiten Klassen Regeln zur Groß-/Kleinschreibung (4 Unterrichtsstunden) und daneben zur Strategie der Ableitung von Schreibungen (5 Unterrichtsstunden). Im Vergleich zur Kontrollgruppe mit gleicher Eingangsleistung war im ersten Fall die Fehlerzahl um fast 30 %, im zweiten um 15 % niedriger. Deutlich höher waren die Zugewinne, die Scheerer-Neumann (1979, 115) in einem kombinierten Regel-, Strategie- und Wortschatz-Training besonders schwacher Rechtschreiber/innen erzielte: während sich die Kontrollgruppe nur um 20 % verbesserte, sank die Fehlerquote – zumindest im Bereich der Groß-/Kleinschreibung – in der Übungsgruppe um 75 %, (vgl. zur Übung der Doppelkonsonanz nach Kurzvokal auch die positiven Befunde bei Schulte-Körne u. a. 2003).

Ein solches explizites Regeltraining ist wegen der komplexen Struktur der deutschen Orthographie aber nur für wenige Bereiche möglich. Zudem stellt sich die Frage, ob bzw. in welchen Formen es überhaupt nötig ist (s. dazu unten). So hat Afflerbach (1997, 67) gezeigt, dass Schüler/innen am Ende der Grundschule 87 % der Kommata richtig setzen – ohne dass deren Systematik bis dahin Gegenstand des Unterrichts gewesen wäre – und dass sie andererseits danach von systematischer Instruktion zur Zeichensetzung nicht erkennbar mehr profitierten (vgl. zu den Schwächen von Studierenden in diesem Bereich: Schübel/Piesnack 2005).

**Für die Leistungsbeurteilung bedeutsam** ist der Befund, dass sich die Fehlerquote von guten Rechtschreiber/innen zwischen erster und zweiter Hälfte von Diktaten kaum verändert, bei schwachen Rechtschreiber/innen aber im zweiten Teil erheblich zunimmt (Schneider 1985; Brügelmann 1994b). Dies ist ein wichtiges Zusatzargument gegen Diktate als Leistungsnachweis und für weniger stress-erzeugende Formen des Leistungsnachweises, z. B. »Pässe« mit definierten Anforderungen, denen sich die Kinder zu unterschiedlichen Zeiten (und ggf. mehrfach) stellen können, oder offenere Aufgaben, die den Kindern ermöglichen, ihre aktuellen Kompetenzen zu zeigen (z. B. »Schreibt möglichst viele Wörter mit einem kurzen Selbstlaut auf – und ordnet sie nach unterschiedlichen Schreibweisen«).

## Sprachsystematische Konzepte

In diesem Bereich sind verschiedene Ansätze zu unterscheiden. Neben dem Morphem (vgl. zur systematischen Nutzung eher mit älteren Schüler/innen: Schubenz 1981) spielt im Anfangsunterricht die Silbe als Einheit unterhalb der Wortebene eine wichtige Rolle (für das Lesenlernen schon belegt durch Scheerer-Neumann 1981 und Adams u. a. 1980). In dieser Form (als *Sprechsilbe*) ist sie den meisten Kindern schon aus Liedern und Reimen vertraut, sodass sich viele auch in ersten Schreibversuchen auf diese Artikulationseinheit beziehen – vor allem in romanischen Sprachen mit einfachen Konsonant-Vokal-Folgen (vgl. die Beispiele in Ferreiro/Teberosky 1982, 197ff.). Davon zu unterscheiden sind explizit sprachanalytische Ansätze mit Bezug sich auf die *Rechtschreibsilbe* (z. B. Röber 2006 im Anschluss an Maas). Zu diesem Ansatz hat Weinhold (2006; 2009) einen mehrjährigen Längsschnitt durchgeführt. In ihrem Methodenvergleich zu Schreibungen vorgegebener Wörter von Klasse 1 bis 4 fand sie zum einen unterschiedliche Entwicklungsrhythmen der Kinder, die nach »Lesen durch Schreiben«, mit einer Fibel oder nach dem silbenanalytischen Ansatz unterrichtet worden waren (2006b): Im Leistungsvergleich lag mal der eine Ansatz vorne, mal ein anderer – außerdem gab es Unterschiede je nach Schwerpunkt der eingesetzten Tests und zum Teil auch in den Fehlerarten. In ihrem Bericht nach Abschluss von Klasse 4 resümiert Weinhold (2009, 71), dass sich bis zum Ende der Grundschulzeit Fibel- und Silbenklassen im Leistungsniveau weitgehend angleichen. Für die Rechtschreibung in freien Texten kommt Fay (2010, 165; s. auch 2012) zu ähnlich geringen Unterschieden.

In Feinanalysen von Fehlern konnten allerdings – bezogen auf einzelne Kinder(gruppen) – durchaus konzeptspezifische Methodeneffekte festgestellt werden (vgl. Weinhold 2010; Heitmann 2012). Diese sprechen dafür, dass für einen Teil der Kinder und in bestimmten Phasen ein silbenanalytisches Vorgehen förderlich sein kann. Auf Gruppenebenen war – wie auch in

anderen Konzeptvergleichen – die methodeninterne Streuung der Wirkungen allerdings so groß, dass lehrer- und schülerspezifische Einflüsse höher eingeschätzt werden müssen (Heitmann 2012, 90).

Zum sprachsystematischen Ansatz von Hinney (1997, im Anschluss an Eisenberg) berichten Blatt/Jarsinski (2009, 110f.) positive Befunde sowohl im Anfangsunterricht als auch in der Förderung von Schüler/innen mit besonderen Schwierigkeiten. Auch wenn die Stichproben keine Generalisierungen erlauben, belegt die vergleichende Fallstudie von zwei ersten Klassen das Potenzial des Ansatzes in ganz verschiedenen Leistungsdimensionen. Und für ältere Schüler/innen erbrachte eine Zusatzförderung von 600 Fünftklässler/innen mit besonderen Schwierigkeiten über ein Schuljahr hinweg Lernzuwächse, die der durchschnittlichen Lernentwicklung über zwei Schuljahre hinweg entsprachen.

## Aneignungstheoretische Ansätze

Diese Konzepte fokussieren die – auch beim Erst- und Fremdsprachenlernen beobachtete – implizite Musterbildung (vgl. Neuweg 2015) und sind insofern charakteristisch für den Spracherfahrungsansatz, der auf ein Rechtschreiblernen »im Gebrauch« (Kochan 1995) setzt. Ausgangspunkt ist die Aneignung der alphabetischen Strategie durch das lautorientierte Verschriften von Wörtern. Auswertungen der angelsächsischen Forschung betonen seit längerem die produktive Funktion des lautorientierten Spontanschreibens (»invented spelling«) für die Anfangsphase des Schriftspracherwerbs – u. a. als Medium beiläufiger Förderung der phonologischen Bewusstheit (Treimann 1985; Mann et al. 1987; Richgels 1995; 2001; Torgesen/Davis 1996; National Reading Panel 2000; Ehri/Roberts 2006). Ähnliche Befunde berichten aus dem deutschsprachigen Raum Brügelmann u. a. (1994b+c) und Weinhold (2006b).

Brinkmann u. a. (2006) konnten bei Kindern mit gering entwickelten Voraussetzungen schon durch kurze Intensivphasen freien Schreibens deutliche Fortschritte in der alphabetischen Strategie erzielen. Nach Neubauer/Kirchner (2013) sind diese Voraussetzung für erfolgversprechende orthographische Übungen. Das besondere Potenzial dieses Ansatzes zeigen einerseits sehr erfolgreiche Klassen, die nach dem Reichen-Konzept unterrichtet wurden (vgl. etwa Brügelmann u. a. 1994b; Hüttis-Graff 1997), auch für die Förderung von Kindern mit besonderen Schwierigkeiten (vgl. Peschel 2003 und i. d. B.), andererseits auffallend schwache Ergebnisse einiger Klassen in anderen Studien (vgl. die kritischen Hinweise von Metzke 2008). Die kleinen Stichproben schränken aber auch hier die Möglichkeit zur Generalisierung ein.

Ein interessantes Teilergebnis erbrachte die Untersuchung von Nickel (1979) zur **Art der Rückmeldung und Korrektur von Prüfdiktaten**. Bei unmittelbarer anschließender Selbstkorrektur durch die Schüler/innen war die Fehlerquote in einer zweiten Überprüfung um gut 20% niedriger als bei einer um zwei Tage verzögerten Rückmeldung durch die Lehrperson – und das, obwohl die Schüler/innen rund ein Viertel ihrer Fehler gar nicht erkannten. Brinkmann (1993) verglich in einer Kindergruppe im Wechsel drei verschiedene Korrekturformen: Selbstkontrolle (wie bei Nickel); dreimaliges Berichtigten angestrichener Fehler; Diskussion der von den Kindern selbst gemeldeten Rechtschreibschwierigkeiten in der Klasse. Nach vier Korrekturdurchgängen sank die Fehlerzahl in allen drei Varianten auf etwa die Hälfte. Bei verschiedenen Kindern erwiesen sich die drei Korrekturformen allerdings als sehr unterschiedlich effektiv. Brinkmann empfiehlt deshalb, diese von den Kindern bewusst ausprobieren zu lassen und sie darin zu unterstützen, möglichst selbstständig in der Korrektur der eigenen Texte zu werden. So wird es auch als grundlegendes Ziel des Rechtschreibunterrichts in den Standards der KMK (2004) gefordert.

Die Sorge allerdings, dass Kinder durch das lautorientierte Konstruieren von Wörtern Fehlstrategien entwickeln, haben Längsschnittbefunde des Bremer Projekts »Kinder auf dem Weg zur Schrift« schon in den 1980er-Jahren widerlegt: »Je mehr komplette Lautumschriften ein Kind im Januar schafft (auch wenn sie Rechtschreibfehler enthalten), umso höher ist sein RS-Entwicklungsstand in Juni [des 1. Schuljahres] (.63). Dieser Prädiktor ist mindestens so gut wie die Zahl der im Januar [von den gleichen Kindern] völlig richtig geschriebenen Wörter (.60) oder die Zahl der zu diesem Zeitpunkt verwendeten RS-Muster (.52). In der Tendenz schreiben diese Kinder im Juni auch schon mehr Wörter orthographisch korrekt (.43)« (Brügelmann u. a. 1987/88, 3.11 E, sowie 1999; vgl. analog selbst für die weniger durchsichtige englisch Orthographie: National Early Literacy Panel 2008, 59). Eichler/ Thomé (1995) haben gezeigt, dass sogar falsche Schreibungen ein höheres orthographisches Niveau signalisieren können als Richtigschreibungen, die durch möglicherweise »inkorrekte oder unzureichende »Innere Regeln« produziert« (a. a. O., 41) wurden.

Für die anschließende Phase der individuellen Regelbildung (s. die Darstellung bei Brinkman i. d. B., Teil II) gibt es kaum aussagekräftige Studien. Das hängt zum einen damit zusammen, dass sehr unterschiedliche Konzepte unter Etiketten wie »lernwegorientierter Ansatz« untersucht worden sind. So evaluierten etwa Deimel/ Schulte-Körne (2005; 2006) die »Rechtschreibwerkstatt« von Sommer-Stumpfenhorst – mit eher schlechteren Leistungen (unter allerdings wenig kontrollierten Rahmenbedingungen). Dagegen untersuchten Brügelmann u. a. (1994a, 131; 1994b) die ihrerseits sehr unterschiedlichen Fortsetzungen von »Lesen durch Schreiben« bzw.

einen von Lehrer/inne/ n selbst als mehr oder weniger offen klassifizierten Unterricht (vgl. zur Problematik solcher Selbsteinschätzungen: Hanke 2001, 384). Sie stellten – ähnlich wie später Weinhold (2009) – unterschiedliche Entwicklungsrhythmen für die einzelnen Ansätze fest; am Ende der Grundschulzeit waren für die Rechtschreibung in freien Texten aber keine bedeutsamen Leistungsdifferenzen mehr erkennbar.

Insgesamt sind die – teilweise sogar intern – widersprüchlichen Befunde aus den verfügbaren Vergleichsstudien nur schwer auf einen Nenner zu bringen (vgl. einerseits die eher kritische Einschätzung bei Metzke 2008 und die eher positive bei Brügelmann/ Brinkmann 2013). Funke (2014) hat in seiner Metaanalyse ausgewählter Studien Leistungen aus »Lesen durch Schreiben«-Klassen mit denen aus Fibelklassen verglichen. Über eine statistische Verrechnung der berichteten Effekte kommt er zu dem Ergebnis, »dass eine schlechtere Rechtschreibleistung von Lesen-durch-Schreiben-Klassen gegenüber Fibelklassen in Klassenstufe 2–4 bei gleichen Eingangsvoraussetzungen derzeit nicht belegt ist« (a. a. O., 24). Der Grund ist, »dass die Aussagekraft der vorliegenden empirischen Studien zu Lesen durch Schreiben durch methodologische Beschränkungen begrenzt wird« (a. a. O., 37). *Überhaupt nicht berücksichtigt sind in den Studien Effekte eines frühen Schreibens eigener Texte auf die Schreibmotivation der Kinder und für die Entwicklung ihrer Fähigkeit, inhaltlich und sprachlich anspruchsvolle Texte zu verfassen.*

## Vorläufige Bilanz

Gemessen an den Leistungen in standardisierten Tests erweist sich keine Methode/kein Ansatz als grundsätzlich überlegen. Die methodeninterne Streuung von »Erfolgen« ist wesentlich höher als die Differenzen zwischen den Methoden. Das heißt: Andere Faktoren wie die Lehrkompetenz (vgl. Fay 2010, 165) oder die Umwelt der Lerner/innen (Roos/ Schöler 2009, 235, 247) haben eine deutlich höhere Bedeutung. Dass die (Vergleichs-)Studien wenig zu Grundsatzentscheidungen zwischen den »großen« Ansätzen beitragen (können), liegt auch an einer Reihe von forschungsmethodischen Einschränkungen, die generell für Studien im pädagogischen Feld gelten:

- Die Definitionen der Ansätze variieren zwischen den Untersuchungen.
- Die Konzepttreue der Umsetzung ist meist nicht oder nur oberflächlich erfasst.
- Die Stichproben sind in der Regel sehr klein (als Einheiten müsste für die meisten Fragestellungen nach Klassen, nicht nach Kindern ausgewertet werden!).
- Zufallsstichproben oder gar Doppelblind-Studien, die den Einfluss anderer Faktoren minimieren, sind bei Untersuchungen im Feld nicht möglich.

Insofern sind es eher die Befunde zu Teilfragen, die in der Praxis weiterhelfen (s. Kästen), z.B. zu Art / Auswahl und Umfang eines Übungswortschatzes (s. o.).

Generell wird in diesen Studien zur Effektivität von Methoden der Einfluss der Lehrperson, der sich in der starken konzept-internen Streuung der Wirkungen zeigt, zu wenig beachtet. Eichler/Brügelmann (2013) ziehen deshalb folgendes Fazit: »Für jede Methode gibt es Beispiele erfolgreichen Unterrichts. Dieser setzt allerdings voraus, dass den Lehrpersonen neben den Potenzialen des jeweiligen Ansatzes auch deren spezifische Risiken bewusst sind – und dass sie über das didaktisch-methodische Repertoire verfügen, diese aufzufangen.«

Es muss aber auch bedacht werden, wie unterschiedlich einzelne Kinder von denselben Methoden profitieren – eine Einsicht, die eher aus Fallanalysen als aus Großstudien zu gewinnen ist (vgl. etwa Bohnenkamp 1994 zu Rechtschreibübungen am Computer vs. Papier-und-Bleistift-Aufgaben und Brinkmann 1993 zur Wirkung verschiedener Korrekturformen). Auch diese Befunde verweisen auf die Notwendigkeit eines breiteren methodischen Repertoires – und der Kompetenz der Lehrperson, dieses situationsgerecht einzusetzen. Für die zukünftige Forschung folgt daraus, dass wir differenziertere Beobachtungs- und Interviewstudien brauchen, in denen der Unterricht von Lehrer/innen, die mit möglichst unterschiedlichen Ansätzen erfolgreich sind, dokumentiert und analysiert wird. Auf diesem Wege könnte man deren jeweils besonderen Potenziale und Risiken aufweisen – und nach Bedingungen suchen, unter denen erstere besonders gut entfaltet und letztere möglichst weitgehend vermieden werden. Das erscheint für die Entwicklung der Praxis hilfreicher als die weitere Suche nach einer »im Durchschnitt« erfolgreicherer Methode.

**Auszug aus: Brinkmann, E. (Hrsg.) (2015):  
Rechtschreiben in der Diskussion –  
Schriftspracherwerb und Rechtschreibunterricht.  
Beiträge zur Reform der Grundschule, Bd. 140.  
Grundschulverband: Frankfurt.**