



Marcus Hasselhorn · Wolfgang Schneider
Ulrich Trautwein (Hrsg.)

Lernverlaufs- diagnostik

Tests und Trends

Jahrbuch der pädagogisch-psychologischen Diagnostik

N. F. Band 12

HOGREFE



Lernverlaufsdagnostik

Jahrbuch der pädagogisch-psychologischen Diagnostik

Tests und Trends

Neue Folge Band 12

Lernverlaufsdagnostik

hrsg. von Prof. Dr. Marcus Hasselhorn, Prof. Dr. Wolfgang Schneider
und Prof. Dr. Ulrich Trautwein

Herausgeber der Reihe:

Prof. Dr. Marcus Hasselhorn, Prof. Dr. Wolfgang Schneider,
Prof. Dr. Ulrich Trautwein

Lernverlaufs- diagnostik

herausgegeben von

Marcus Hasselhorn, Wolfgang Schneider
und Ulrich Trautwein

HOGREFE 

GÖTTINGEN · BERN · WIEN · PARIS · OXFORD · PRAG
TORONTO · BOSTON · AMSTERDAM · KOPENHAGEN
STOCKHOLM · FLORENZ · HELSINKI

Prof. Dr. Marcus Hasselhorn, geb. 1957. 1977–1983 Studium der Psychologie und Pädagogik. 1986 Promotion. 1993 Habilitation. 1993–1997 Professor für Entwicklungspsychologie an der TU Dresden. 1997–2007 Leiter der Abteilung Pädagogische Psychologie und Entwicklungspsychologie an der Universität Göttingen. Seit 2007 Leiter der Arbeitseinheit Bildung und Entwicklung am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt am Main.

Prof. Dr. Wolfgang Schneider; geb. 1950. 1969–1975 Studium der Psychologie, Theologie und Philosophie. 1976–1981 Wissenschaftlicher Mitarbeiter am Psychologischen Institut der Universität Heidelberg. 1979 Promotion. 1981–1982 Visiting Scholar an der Stanford University (USA). 1982–1991 Wissenschaftlicher Mitarbeiter am Max-Planck-Institut für psychologische Forschung in München. 1988 Habilitation. 1990–1991 Vertretung und seit 1991 Inhaber des Lehrstuhls für Pädagogische und Entwicklungspsychologie an der Universität Würzburg.

Prof. Dr. Ulrich Trautwein, geb. 1972. 1992–1999 Studium der Psychologie. 1999 Diplom in Psychologie. 2002 Promotion. 2005 Habilitation. Seit 2008 Universitätsprofessor für Empirische Bildungsforschung an der Universität Tübingen.

© 2014 Hogrefe Verlag GmbH & Co. KG
Göttingen · Bern · Wien · Paris · Oxford · Prag · Toronto · Boston
Amsterdam · Kopenhagen · Stockholm · Florenz · Helsinki
Merkelstraße 3, 37085 Göttingen

<http://www.hogrefe.de>

Aktuelle Informationen · Weitere Titel zum Thema · Ergänzende Materialien

Copyright-Hinweis:

Das E-Book einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.

Der Nutzer verpflichtet sich, die Urheberrechte anzuerkennen und einzuhalten.

Satz: ARThür Grafik-Design & Kunst, Weimar
Format: PDF

ISBN 978-3-8409-2614-3

Nutzungsbedingungen:

Der Erwerber erhält ein einfaches und nicht übertragbares Nutzungsrecht, das ihn zum privaten Gebrauch des E-Books und all der dazugehörigen Dateien berechtigt.

Der Inhalt dieses E-Books darf von dem Kunden vorbehaltlich abweichender zwingender gesetzlicher Regeln weder inhaltlich noch redaktionell verändert werden. Insbesondere darf er Urheberrechtsvermerke, Markenzeichen, digitale Wasserzeichen und andere Rechtsvorbehalte im abgerufenen Inhalt nicht entfernen.

Der Nutzer ist nicht berechtigt, das E-Book – auch nicht auszugsweise – anderen Personen zugänglich zu machen, insbesondere es weiterzuleiten, zu verleihen oder zu vermieten.

Das entgeltliche oder unentgeltliche Einstellen des E-Books ins Internet oder in andere Netzwerke, der Weiterverkauf und/oder jede Art der Nutzung zu kommerziellen Zwecken sind nicht zulässig.

Das Anfertigen von Vervielfältigungen, das Ausdrucken oder Speichern auf anderen Wiedergabegeräten ist nur für den persönlichen Gebrauch gestattet. Dritten darf dadurch kein Zugang ermöglicht werden.

Die Übernahme des gesamten E-Books in eine eigene Print- und/oder Online-Publikation ist nicht gestattet. Die Inhalte des E-Books dürfen nur zu privaten Zwecken und nur auszugsweise kopiert werden.

Diese Bestimmungen gelten gegebenenfalls auch für zum E-Book gehörende Audiodateien.

Anmerkung:

Sofern der Printausgabe eine CD-ROM beigelegt ist, sind die Materialien/Arbeitsblätter, die sich darauf befinden, bereits Bestandteil dieses E-Books.

Inhaltsverzeichnis

| | |
|--|-----|
| Vorwort der Herausgeber | VII |
| Kapitel 1 Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdagnostik <i>Karl Josef Klauer</i> | 1 |
| Kapitel 2 Formative Leistungsdiagnostik in der Sekundarstufe – Grundlegende Fragen, domänenspezifische Verfahren und empirische Befunde <i>Uwe Maier</i> | 19 |
| Kapitel 3 Formative Leistungsbeurteilung im Unterricht: Konzepte, Praxisberichte und ein neues Diagnoseinstrument für das Fach Mathematik <i>Anika Bürgermeister, Eckhard Klieme, Katrin Rakoczy, Birgit Harks und Werner Blum</i> | 41 |
| Kapitel 4 Diagnostik und Prävention von Lernschwierigkeiten im Aptitude Treatment Interaction-(ATI-) und Response to Intervention-(RTI-)Ansatz <i>Yvonne Blumenthal, Kristin Kuhlmann und Bodo Hartke</i> | 61 |
| Kapitel 5 Curriculumbasierte Messverfahren (CBM) als Methode der formativen Leistungsdiagnostik im RTI-Ansatz <i>Stefan Voß und Bodo Hartke</i> | 83 |
| Kapitel 6 Das Rügener Inklusionsmodell (RIM) – RTI in der Praxis <i>Kathrin Mahlau, Yvonne Blumenthal, Kirsten Diehl, Anne Schöning, Simon Sikora, Stefan Voß und Bodo Hartke</i> | 101 |
| Kapitel 7 Fähigkeitsindikatoren Primarschule (FIPS) – Überprüfung des Lern- erfolgs in der ersten Klasse <i>Kerstin Bäuerlein, Frank Niklas und Wolfgang Schneider</i> | 127 |
| Kapitel 8 Lesekompetenzen formativ evaluieren mit dem IEL-1 – Inventar zur Erfassung der Lesekompetenzen von Erstklässlern <i>Kirsten Diehl</i> | 145 |

Kapitel 9**Lernfortschrittsdiagnostik Lesen (LDL) und Verlaufsdagnostik sinn-
erfassenden Lesens (VSL): Zwei Verfahren als Instrumente einer formativ
orientierten Lesediagnostik***Jürgen Walter* 165**Kapitel 10****Lernverlaufsdagnostik Mathematik für zweite bis vierte Klassen (LVD-M)***Alfons M. Strathmann* 203**Kapitel 11****Wirksamkeit formativen Assessments – Evaluation des Ansatzes
der Lernverlaufsdagnostik***Elmar Souvignier, Natalie Förster und Elisabeth Schulte* 221**Kapitel 12****quo: Ein Ansatz internetbasierter Lernverlaufsdagnostik
mit Testkonzepten für Lesen und Mathematik***Elmar Souvignier, Natalie Förster und Martin Salaschek* 239**Kapitel 13****Lernentwicklungsmonitoring mit KEKS***Peter May, Jasmine Bennöhr und Carina Berger* 257**Kapitel 14****Instrumente zur Lernverlaufsmessung: Gütekriterien und Auswertungs-
herausforderungen***Jürgen Wilbert* 281**Autorenverzeichnis** 309

Vorwort der Herausgeber

Seit dem Beginn der Neuen Folge der Reihe „Tests und Trends – Jahrbuch der pädagogisch-psychologischen Diagnostik“ im Jahr 2000 wurde in den nun 14 Jahren ihres Bestehens immer wieder versucht, den neuesten Stand diagnostischer Möglichkeiten in unterschiedlichen Inhaltsbereichen schulischen Lernens zu dokumentieren. In den bislang erschienenen Bänden der Reihe wurden praktisch relevante Forschungsansätze und -ergebnisse zu spezifischen Lernleistungen, spezifischen Lernvoraussetzungen sowie zu lernbegleitenden Fähigkeiten, Funktionen und Dispositionen unter diagnostischem Blickwinkel berichtet. Im Mittelpunkt der Betrachtung standen dabei meist standardisierte Testverfahren zur Erfassung von Kompetenzen in ausgewählten Inhaltsbereichen, die als „Status-tests“ das Leistungsvermögen von Schülerinnen und Schülern zu einem bestimmten Zeitpunkt erfassen. Diese Art von Bestandsaufnahme wird auch als „summativ Diagnostik“ charakterisiert und bezeichnet in der Regel die Evaluation des Ergebnisses eines langfristigen Lernvorgangs.

Der hier vorliegende 12. Band der Neuen Folge beschäftigt sich im Unterschied dazu mit Möglichkeiten der „formativen Leistungsdiagnostik“, die schon seit mehreren Jahrzehnten in den USA relativ populär ist und in den letzten Jahren auch verstärkt im deutschsprachigen Raum Beachtung gefunden hat. Wenn auch die Definition dieses Konzepts in der Literatur nicht ganz eindeutig ausfällt, so ist hier meist eine systematische *Lernverlaufsdagnostik* gemeint, die Informationen über die Entwicklung von Schülerleistungen gibt und von der Grundannahme geleitet wird, dass die beständige Rückmeldung von Leistungsveränderungen an Schülerinnen und Schüler sowie Lehrerinnen und Lehrer zu insgesamt besseren Ergebnissen führt.

Der vorliegende Band gibt den Stand der neueren deutschsprachigen Entwicklungsarbeit im Bereich der *formativen* Diagnostik und Evaluation wieder. Das einleitende Kapitel von *Karl Josef Klauer* erläutert die begriffliche wie auch methodische Problematik und führt in die Konzepte der curriculumsbasierten Messung und des Ansatzes der „Response to Intervention“ (RTI) ein, die in den Kapiteln des Bandes eine wichtige Rolle spielen. Das sich anschließende Kapitel von *Uwe Maier* gibt einen Überblick über theoretische Grundannahmen und Methoden der formativen Leistungsdiagnostik in der Sekundarstufe und illustriert Anwendungsmöglichkeiten im mathematisch-naturwissenschaftlichen wie auch im sprachlichen Unterricht, die auch Ansätze computergestützter formativer Leistungsdiagnostik mit einschließen. In ähnlicher Weise überblicksorientiert ist das Kapitel zur formativen Leistungsbeurteilung von *Anika Bürgermeister, Eckhard Klieme, Katrin Rakoczy, Birgit Harks und Werner Blum*, das ebenfalls den Unterricht der Sekundarstufe in den Blick nimmt und die Möglichkeiten dieses Ansatzes im Rahmen eines DFG-geförderten Projekts zur Leistungsbeurteilung im deutschen Mathematikunterricht illustriert.

Der Beitrag von *Yvonne Blumenthal, Kristin Kuhlmann und Bodo Hartke* kontrastiert den klassischen „aptitude-treatment-interaction“-Ansatz (ATI) mit dem oben erwähnten RTI-Ansatz, um die Möglichkeiten unterschiedlicher Lösungsansätze für das Problem der optimalen Anpassung von Unterrichtsmethoden an unterschiedliche Lernvoraussetzungen von Schülern darzustellen. Das übergeordnete Ziel des RTI-Ansatzes, dessen Vorteil in einem gestuften Fördersystem liegt, wird in der Früherkennung und Prävention von Lernschwierigkeiten und Lernstörungen gesehen, was als sinnvolle Alternative zum ATI-Ansatz gelten kann. Das sich anschließende Kapitel von *Stefan Voß und Bodo Hartke* bezieht sich ebenfalls auf den RTI-Ansatz und verdeutlicht im Rahmen eines Literaturüberblicks, dass hier curriculumbasierte Messverfahren (CBM) für die Erfassung des Lernverlaufs in unterschiedlichen Inhaltsbereichen von zentraler Bedeutung sind. Ein aktuelles Anwendungsbeispiel für den Einsatz von RTI bietet der Beitrag von *Kathrin Mahlau, Yvonne Blumenthal, Kirsten Diehl, Anne Schöning, Simon Sikora, Stefan Voß und Bodo Hartke*, der das „Rügener Inklusionsmodell“ beschreibt und aufzeigt, wie Mehrebenenprävention funktionieren kann, wenn je nach Ausmaß des Lerndefizits in den Bereichen Deutsch und Mathematik, aber auch im Bereich emotionaler und sozialer Entwicklung unterschiedliche Förderebenen (regulärer Unterricht, Kleingruppenförderung, individuelle Förderung) aktiviert werden.

Kerstin Bäuerlein, Frank Niklas und Wolfgang Schneider stellen mit dem Verfahren „Fähigkeitsindikatoren Primarschule“ (FIPS) ein computergestütztes Diagnoseinstrument für den Einsatz in der ersten Schulklasse vor. Über den Vergleich der Lernausgangslage in den Bereichen Wortschatz, Lautbewusstheit, Lesen und Mathematik zu Beginn der Schulzeit mit den im Verlauf des ersten Schuljahrs erzielten Lernfortschritten in diesen Bereichen können wichtige Informationen für Lehrkräfte bereitgestellt werden, die adaptive Fördermaßnahmen zu einem frühen Zeitpunkt ermöglichen. *Kirsten Diehl* stellt mit dem Inventar zur Erfassung der Lesekompetenzen von Erstklässlern (IEL-1) ein formatives Evaluationsverfahren zur Erfassung der Lernfortschritte im Lesen vor, das ähnlich wie FIPS die Schuleingangsphase im Blick hat, dabei aber auf den Leselernvorgang fokussiert und in der formativen Diagnostik so kleinschrittig angelegt ist, dass Rückstände im Leselernprozess frühzeitig erkannt werden können. Möglichkeiten der Lernfortschrittsdiagnostik und der Verlaufsdagnostik im Bereich Lesen werden auch umfassend im Kapitel von *Jürgen Walter* am Beispiel von zwei diagnostischen Verfahren beschrieben, die sorgfältig konzipiert und evaluiert wurden und im Grundschulbereich wie auch in der Sekundarstufe einsetzbar sind. Ermutigende Befunde zur Lernverlaufsdagnostik Mathematik für die zweite bis vierte Klassenstufe der Grundschule liefert der Beitrag von *Alfons M. Strathmann*, in dem ein Verfahren vorgestellt wird, dass mittels CD die Generierung von immer neuen Einzeltests möglich macht. Lernverlaufskurven lassen sich damit problemlos für jedes Kind einer Klasse wie auch die Klasse als Ganzes erstellen.

Zwei Beiträge der Münsteraner Arbeitsgruppe um Elmar Souvignier setzen sich mit der Wirksamkeit formativer Diagnostik und der Möglichkeit von internetba-

sierten Einsätzen auseinander. *Elmar Souvignier, Natalie Förster* und *Elisabeth Schulte* gehen davon aus, dass idealerweise diagnostische Verfahren mit hoher psychometrischer Güte wiederholt zum Einsatz kommen müssen, um die Wirksamkeit des Verfahrens zu optimieren. Sie stellen mögliche Untersuchungsdesigns bei der Wirksamkeitsprüfung formativer Assessments vor und präsentieren Befunde ihres Forschungsprogramms zur Evaluation formativer Diagnostik. Das sich anschließende Kapitel von *Elmar Souvignier, Natalie Förster* und *Martin Salaschek* geht weiter ins Detail und illustriert die Möglichkeiten internetbasierter Lernverlaufdiagnostik in den Bereichen Lesen und Mathematik am Beispiel von „quop“, einem internetbasierten System, das eine ökonomische Durchführung der formativen Diagnostik im Regelunterricht der Grundschule sowie eine automatisierte Auswertung und Dokumentation der Ergebnisse ermöglicht.

Der Beitrag von *Peter May, Jasmine Bennöhr* und *Carina Berger* behandelt die Hamburger Testserie „KEKS“, ein Verfahren zur „Kompetenzerfassung in Kindergarten und Schule“, das Kernkompetenzen in den Bereichen Deutsch, Mathematik, Englisch und – bei Kindern mit Migrationshintergrund – unterschiedlichen Herkunftssprachen erfasst. KEKS erfüllt die üblichen Gütekriterien und ist sowohl bei Kindern als auch bei Jugendlichen einsetzbar, umfasst dabei den Altersbereich ab 4 bis ca. 16 Jahre und kann demnach die individuelle Lernentwicklung über einen größeren Zeitraum abbilden. Im abschließenden Kapitel von *Jürgen Wilbert* geht es um methodische Herausforderungen und Besonderheiten der Lernverlaufdiagnostik und die Illustration messtheoretischer Voraussetzungen, die bei Instrumenten zur Lernverlaufdiagnostik erfüllt sein sollten, wenn sie sich in Forschung und Praxis der empirischen Bildungsforschung, der Sonderpädagogik und der Pädagogischen Psychologie fest etablieren wollen.

Insgesamt gesehen bietet der vorliegende Band einen umfassenden Überblick über neuere Entwicklungen im Bereich der Lernverlaufdiagnostik, die das Potenzial dieses diagnostischen Ansatzes für die Erfassung der Leistungsentwicklung in unterschiedlichen schulischen Inhaltsbereichen und für unterschiedliche Altersgruppen illustrieren, gleichzeitig aber auch die Voraussetzungen für einen effektiven Einsatz deutlich machen.

Frankfurt, Würzburg, Tübingen, im Dezember 2013

Marcus Hasselhorn
Wolfgang Schneider
und Ulrich Trautwein

Kapitel 1

Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdagnostik

Karl Josef Klauer

Zusammenfassung

Die Unterscheidung von formativer und summativer Diagnostik wurde 1967 von Scriven eingeführt. Formative Diagnostik bezog sich auf wiederholte Messungen im Laufe eines Lernprozesses, summative Diagnostik auf die Messung des am Ende erreichten Leistungsstandes. Allerdings erhielt formative Diagnostik in den USA immer wieder neue Interpretationen, was zu vielfach beklagter Konfusion führte. Dennoch resultierten Ansätze, die sich als zukunftsfruchtig erweisen sollten, so etwa das „Curriculum Based Measurement“ oder die „Response to Intervention“. Eine deutsche Entwicklung formativer Evaluation stellt die Lernverlaufsdagnostik dar. Die Lernverlaufsdagnostik erfordert (1) eine klare Definition der wiederholt zu erfassenden Kompetenz, (2) homogen schwierige, (3) änderungssensible Tests und (4) unter Umständen eine Abkehr von der klassischen Testtheorie.

1.1 Historische Entwicklung in den USA

1.1.1 Unterscheidung formativer und summativer Diagnostik

Die Unterscheidung von formativer und summativer Diagnostik geht auf Scriven (1967) zurück. Er bezeichnete formative Evaluation als die in Abständen wiederholte Evaluation eines sich im Vollzug befindlichen Lernprozesses und im Gegensatz dazu summative Evaluation als Evaluation des Ergebnisses am Schluss einer solchen Lernperiode oder eines Lernprogramms. Bloom (1969) griff die Unterscheidung auf und modifizierte sie deutlich. Formative Diagnostik sollte danach nicht nur den Lernfortschritt im Laufe des Prozesses *dokumentieren*, sondern zugleich auch den Prozess *fördern*, das Lernen *verbessern*, und zwar durch *Rückmeldung* der Lernergebnisse an die Lernenden einerseits und an die Lehrenden andererseits. In ihrem auch in Deutschland weit verbreiteten Handbuch sorgten Bloom, Hastings und Madaus (1971) für die allgemeine Verbreitung dieser Unterscheidung von formativer und summativer Evaluation und für die von Bloom vorgeschlagene Interpretation.

Dabei ist die Entwicklung aber nicht stehen geblieben. In der Folge wurde die Unterscheidung Allgemeingut, wobei allerdings viele Autoren die Begriffe mehr oder minder nach eigenem Gutdünken interpretierten. Tatsächlich wurde sehr bald schon eine große Konfusion um die Unterscheidung von „formativ“ und „summativ“ festgestellt, weil sie zwar allgemein akzeptiert und gebraucht, aber vielfach beliebig umgedeutet wurde (Misanchuk, 1978). Das betraf insbesondere die formative Evaluation, die ja auch das eigentlich Neue darstellte, denn die summative Evaluation bezieht sich im Wesentlichen auf die gängigen Schulleistungstests, wie sie etwa am Ende eines Schuljahres eingesetzt werden.

In besonders einflussreichen Abhandlungen hatten Black und Wiliam (1998; 2003) formative Evaluation im Anschluss an Bloom nicht als ein Verfahren definiert, das durch bestimmte *Eigenschaften* gekennzeichnet sei, sondern durch den *Zweck*, dem es dient, nämlich durch die Lernförderung. Ihr wichtigstes Argument bezogen sie aus einer Metaanalyse, die die bekannten Sonderpädagogen Fuchs und Fuchs (1986) veröffentlicht hatten und die glaubten nachzuweisen, dass formative Evaluation das Lernen in erheblichem AusmaÙe voranbringt. Black und Wiliam hielten deshalb für erwiesen, dass formative Evaluation das Lernen fördert. Inzwischen gibt es aber erhebliche und gut begründete Zweifel an der Qualität dieser Metaanalyse und an den daraus gezogenen Schlussfolgerungen, wie Dunn und Mulvenon (2009) in einer elektronischen Zeitschrift differenziert belegen. Allerdings gilt es, hier zu unterscheiden: Tatsächlich gibt es gut begründete Nachweise, dass die regelmäßige Leistungsmessung mit entsprechenden Rückmeldungen die Leistung deutlich verbessern kann, nur ist dies eben nicht unter allen Bedingungen der Fall.

Die allgemein große Zustimmung, die die formative Diagnostik inzwischen erhalten hatte und noch erhält, führte immer mehr Autoren dazu, eigene Konzepte „formativer“, also lernfördernder Evaluation zu entwickeln. Das sollen einige Beispiele illustrieren. Beer und Bloomer (1986) traten für eine Kombination formativer *und* summativer Evaluation ein, wobei die formative Evaluation weitgehend qualitative Methoden verwenden sollte. Oder für Sadler (1989) ging es in der formativen Evaluation primär darum, das „Feedback“ auf differenziertere Weise zu vermitteln. Ferner definierten Nicol und MacFarlane-Dick (2006) formative Evaluation durch die Optimierung des *selbstregulierten* Lernens. Oder Garrison und Ehringhaus (2007) stellten fest, dass formatives Assessment durch aktive Einbeziehung der *Lernenden*, durch *Praxis* und noch durch manches andere gekennzeichnet sei. Am Beispiel der Fahrschule wollten sie klar machen, was sie meinten: Die Fahrprüfung sei summativ, aber die Fahrübungen vorher seien formativ. Winger (2005) schlug sogar vor, formativ-summative Tests einzusetzen, also beide Varianten miteinander kombiniert zu verwenden, die von anderen Autoren als gegensätzlich eingeschätzt worden waren. Misanchuk (1978) war aber Jahre zuvor schon einen Schritt weiter gegangen und hatte eine dritte Variante eingeführt: Formative, summative und konfirmative Evaluation. Wesentlich später waren Dunn und Mulvenon (2009) in der Lage, noch weiter zu diffe-

renzieren und genau vier Varianten einzuführen. Sie unterschieden nun formatives Assessment von summativem Assessment und analog formative Evaluation von summativer Evaluation. Für die Autoren war die nahezu chaotische Konfusion, die sich im englischsprachigen Raum um formativ und summativ sowie um Evaluation und Assessment entwickelt hatte, der Anlass, um gedankliche Ordnung zu stiften. Ob sie damit bei ihren Kollegen Erfolg haben werden, darf man füglich bezweifeln, denn das terminologische Durcheinander ist bislang nicht wesentlich besser geworden.

Ein Motiv für die *Weiterentwicklung* dieser Testtradition wurde in den USA durch das „No Child Left Behind“-Gesetz von 2002 ausgelöst. Schon im Rahmen der älteren *Accountability*-Politik waren die Schulen gehalten, dafür zu sorgen, dass alle Kinder angemessen gefördert würden. Das neue Gesetz von 2002 machte die Schulen nun aber unmittelbar *verantwortlich* für etwaige Misserfolge von Kindern. Seither sind die Schulen gehalten, dass 95 % der Schülerinnen und Schüler die Vorgaben des Bundeslandes im Lesen und in Mathematik erreichen, ab dem Schuljahr 2013/2014 sollen sogar 100 % der Kinder die Norm erfüllen. Das Gesetz belegte die Schulen mit entsprechend harten Sanktionen, die von der Entlassung von Lehrkräften bis zur Schließung ganzer Schulen reichen können (Deno, 2003a). Diese Gesetzgebung führte zu einer Fülle von Aktivitäten, um Leistungsschwächen frühzeitig zu erkennen und um das Lernen entschieden zu fördern. Eine wirklich neue und folgenreiche Entwicklung stammt aus dem Bereich der Sonderpädagogik, das „curriculum-based measurement“, welches allerdings schon weit vor dem Jahre 2002 einsetzte, dann aber durch das Gesetz starken Auftrieb erhielt.

1.1.2 *Curriculum-Based Measurement (CBM)*

Der Ansatz wurde von Stanley Deno und seinen Mitarbeitern an der Universität Minnesota seit 1972 entwickelt (vgl. auch Deno, 1985, 2003a, b). Nicht zufällig entstand das Konzept in der Sonderpädagogik, waren doch leistungsschwache Kinder auch schon von der älteren *Accountability*-Politik in besonderem Maße betroffen und Lehrer entsprechend motiviert, solche Kinder frühzeitig zu erkennen und entsprechend zu fördern. Denos Begriff lässt sich mit curriculumbasierter Messung leicht ins Deutsche übertragen, würde dann aber falsche Assoziationen auslösen. Tatsächlich sind die deutschen Schulleistungstests typischerweise an den Lehrplänen soweit wie möglich orientiert, und es ist in aller Regel möglich, mithilfe eines Schulleistungstests summativ festzustellen, ob ein Kind die Leistungen erbringt, die für das entsprechende Schuljahr vorgegeben sind. In dem Sinne könnte man deutsche Schulleistungstests durchaus als curriculumbasiert bezeichnen.

Genau solche Leistungsmessungen sind aber nicht gemeint mit „Curriculum-Based Measurement“. Hier geht es gerade nicht um den Einsatz der üblichen