

Empirische Sonderpädagogik, 2017, Nr. 2, S. 116-122
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

Kurzbeitrag **Normierung von Verfahren zur Lernverlaufsdagnostik**

Natalie Förster, Jörg-Tobias Kuhn & Elmar Souvignier

Westfälische Wilhelms-Universität Münster

Zusammenfassung

In diesem Kurzbeitrag schildern wir Herausforderungen bei der Normierung von Verfahren zur Lernverlaufsdagnostik, die sich bei der Statusdiagnostik in dieser Form nicht stellen. Diese betreffen insbesondere die Frage, ob Normen für regulären Unterricht oder intensive Förderung benötigt werden, aber auch die Unterschiedlichkeit von Lernzuwächsen in Abhängigkeit von der erfassten Kompetenz, des verwendeten Messverfahrens, des Untersuchungszeitraums und bestimmter Schülermerkmale. Darüber hinaus weisen Lernverläufe im Unterschied zu einmaligen Testungen die statistische Besonderheit auf, dass die Größe der Vertrauensintervalle für den Lernzuwachs von der Anzahl der verfügbaren Messungen abhängt. Basierend auf einer Analyse dieser Herausforderungen schlagen wir Designmerkmale und Analyseschritte bei der Normierung von Verfahren zur Lernverlaufsdagnostik vor.

Schlüsselwörter: Lernverlaufsdagnostik, Normierung, Response to Intervention, Curriculumbasiertes Messen

Establishing standard measures of growth in learning progress assessment

Abstract

The standardization of growth in learning progress assessment faces challenges that go beyond those of typical achievement tests. These challenges concern the question whether norm values refer to business-as-usual instruction or effective treatments, and the variability of learning growth as a function of the respective competence, assessment, seasonal effects, and student characteristics. Moreover, in contrast to confidence intervals of single assessments, confidence intervals of slopes are statistically specific in that they depend on the number of assessments. Based on these challenges, we propose design features and steps for the analysis when establishing standard measures of growth for learning progress assessment.

Keywords: Learning progress assessment, Standardization, Response to intervention, Curriculum-based Measurement

Mit Lernverlaufsdagnostik werden diagnostische Verfahren bezeichnet, die im Sinne eines formativen Assessments Lernverläufe von Schülerinnen und Schülern über kurze,

wiederholte Testungen erfassen, um diagnosebasierte pädagogische Entscheidungen zu ermöglichen und Fördermaßnahmen zu evaluieren. Derartige Verfahren kommen

dabei in sehr unterschiedlichen Bereichen zum Einsatz. So finden sie u.a. Verwendung im Rahmen von multiplen Einzelfallstudien (Grosche, 2011) oder im Response to Intervention-Paradigma (Voß et. al, 2014), haben sich aber auch im regulären Schulunterricht bewährt (Förster & Souvignier, 2014, 2015). Unabhängig vom konkreten Anwendungsrahmen besteht das Ziel der Lernverlaufsdiagnostik immer darin, eine Aussage über die Lernentwicklung zu treffen. Aufgrund der wiederholten Messungen ist dabei grundsätzlich die Anwendung einer individuellen Bezugsnorm möglich, wobei optimaler Weise eine erste Baseline-Messung mit einer späteren Interventionsphase verglichen wird (idiographischer Ansatz). Ein derartiger Vergleich führt jedoch im besten Falle lediglich zu der Aussage, dass die Lernentwicklung unter der Intervention (signifikant) höher ausfällt als zuvor. Eine Aussage darüber, ob es sich um einen außergewöhnlich hohen oder eher um einen vergleichsweise niedrigen Lernzuwachs handelt, erfordert einen weiteren (sozialen) Vergleichsmaßstab in Form einer Norm (nomothetischer Ansatz). Darüber hinaus werden die Lernentwicklungen von Schülerinnen und Schülern häufig mit einer sogenannten Ziellinie verglichen. Hierfür muss aber bekannt sein, welche Lernfortschritte typischerweise zu erwarten sind. Ausgehend von der Überlegung, dass eine Kombination idiographischer und nomothetischer Ansätze insbesondere im Rahmen der Lernverlaufsdiagnostik fruchtbar ist (Asendorpf, 2000), soll dieser Kurzbeitrag Herausforderungen an eine derartige Normierung von Verfahren zur Lernverlaufsdiagnostik skizzieren und mögliche Vorgehensweisen bei der Entwicklung von geeigneten Normwerten vorstellen. Unsere Überlegungen beziehen sich dabei auf Verfahren, die Lernverläufe mittels inhaltlich äquivalenter Paralleltests erheben, da dieser Ansatz den meisten deutschsprachigen Verfahren zugrunde liegt.

Herausforderungen bei der Normierung von Verfahren zur Lernverlaufsdiagnostik

Die zentrale Anforderung bei der Normierung von Tests zur Lernverlaufsdiagnostik besteht darin, dass anhand wiederholter Messungen eine reliable und valide Erfassung von *Veränderung* (und nicht nur des Niveaus wie bei der Statusdiagnostik) erfolgen soll. Entsprechend müssen Normen für die Veränderung einer Kompetenz über die Zeit bereitgestellt werden.

Eine weitere Herausforderung liegt in der Vielfalt der konkreten Verfahren und Anwendungsziele der Lernverlaufsdiagnostik. Dies gilt selbst dann, wenn man die Überlegungen auf Verfahren beschränkt, die auf einem Paralleltestansatz beruhen. Unterschiede beziehen sich dabei insbesondere auf die Dimensionalität des verwendeten Testkonzepts, die Anzahl und Abstände der Messungen, die Testteilnehmer (einzelne Schülerinnen und Schüler bis hin zu ganzen Klassen) und den unterrichtlichen Rahmen (spezifische Förderung vs. regulärer Unterricht). Die Komplexität des Themas steigert sich zudem durch die Unterschiedlichkeit von Lernverläufen. So folgen Leistungsentwicklungen von Schülerinnen und Schülern nicht zwingend linearen Verläufen (Fuchs, Fuchs, Hamlett, Walz & Germann, 1993, Strathmann & Klauer, 2010) und es lassen sich unterschiedliche Verlaufstypen finden (Salaschek, Zeuch & Souvignier, 2014). Darüber hinaus zeigen Studien, dass sich die Lernverläufe abhängig von Untersuchungszeitraum (Christ, Silberglitt, Yeo & Cormier, 2010), Ausgangsniveau (Silberglitt & Hintze, 2007) und anderen Schülermerkmalen (Keller-Margulis, Clemens, Im, Kwok & Booth, 2012) unterscheiden. Bei der Normierung von Lernverlaufsdaten sollte entsprechend geprüft werden, ob Angaben zu durchschnittlichen Veränderungen als verlässlich für alle Schülerinnen und Schüler angenommen werden können.

Ergänzend muss bedacht werden, dass Verfahren zur Lernverlaufsdiagnostik auf-

grund ihres formativen Charakters explizit den Anspruch erheben, diagnostische Informationen für pädagogische Entscheidungen zur Verfügung zu stellen. Daraus ergibt sich die Anforderung, neben der Bereitstellung normativer Verlaufsinformationen zusätzlich Entscheidungsregeln anzugeben, ab wann eine Entwicklung soweit von der Norm abweicht, dass ein Wechsel des Förderansatzes angezeigt ist.

Im Folgenden beschreiben wir zunächst, welche Schlussfolgerungen sich aus diesen Herausforderungen hinsichtlich der Untersuchungsdesigns bei der Normierung von Verfahren zur Lernverlaufdiagnostik ableiten lassen. Im Anschluss stellen wir kurz Schritte und Methoden zur Analyse der Normierungsdaten vor.

Designmerkmale bei der Normierung von Verfahren zur Lernverlaufdiagnostik

Die Empfehlung eines einheitlichen, standardisierten Vorgehens bei der Normierung ist sicherlich nicht möglich. In Abhängigkeit vom konkreten Einsatzzweck des Verfahrens müssen individuelle Anpassungen und Entscheidungen getroffen werden. Folgende Aspekte sollten dabei in die Überlegungen zum Design einer Normierung einfließen.

Repräsentativität der Stichprobe, des Unterrichts und der Testsituation

Grundlegende Voraussetzung für eine geeignete Normierung ist zunächst die präzise Beschreibung der Normierungsstichprobe, der Art des Unterrichts bzw. der Förderung sowie der Testsituation. Es muss gewährleistet sein, dass die drei genannten Merkmale für spätere Vergleiche repräsentativ in dem Sinne sind, dass sie enge Entsprechungen zu der Anwendungssituation aufweisen. Beispielsweise sollte eine Normierung Angaben zu spezifischen Leistungsniveaus der getesteten Schülerinnen und Schüler umfassen, da sich die Lernzuwächse in Abhängig-

keit vom Ausgangsniveau unterscheiden können (Silberglitt & Hintze, 2007). Wenn die durchschnittliche Leistungsentwicklung leistungsschwacher Schülerinnen und Schüler objektiv niedriger läge als bei leistungstärkeren Kindern (Matthäus-Effekt), dann würde eine über alle Niveaus ermittelte mittlere Veränderungsnorm die schwächeren Schüler benachteiligen. Entsprechend bestünde das Risiko, eine – bezogen auf ein spezifisches Ausgangsniveau – effektive Förderung fälschlicherweise als ineffektiv zu interpretieren, nur weil die Veränderung unter der (durchschnittlichen) Norm für alle Schülerinnen und Schüler liegt (vgl. Silberglitt & Hintze, 2007).

Nicht nur die Stichprobe, sondern auch der Unterricht (intensive Förderung vs. regulärer Unterricht) sowie die Testsituation als solche sollte repräsentativ für den späteren Einsatz sein, um zu realistischen Schätzungen der Leistungsentwicklung zu kommen. Dies ist im Rahmen der Lernverlaufdiagnostik insofern besonders wichtig, als die Tests nicht nur sensitiv für Leistungsveränderungen, sondern auch für spezifische Bedingungen der jeweiligen Testsituation (z.B. Ort der Testung, Testleiter) sind, die gerade im Fall der Lernverlaufdiagnostik nur bedingt kontrollierbar sind (Christ, Monaghan, Zopluoglu & Van Norman, 2012).

Ausreichende Anzahl an Messzeitpunkten und ausreichend langer Zeitraum

Untersuchungen zur Reliabilität und Validität von Veränderungsmessungen zeigen, dass deren Güte insbesondere von der Reliabilität der Einzelmessungen sowie der Länge des Zeitraumes, währenddessen die Veränderungen erfasst werden, beeinflusst wird (Christ, Zopluoglu, Monaghan & Van Norman, 2013). Damit Veränderungsmessungen ausreichend reliabel und valide ausfallen können, ist bereits bei wöchentlichen Messungen ein Zeitraum von mindestens sechs Wochen notwendig (Christ et al., 2013)¹. Entsprechend müssen auch Normierungsstudien

für das jeweilige Testverfahren ausreichend viele Messzeitpunkte veranschlagen. Da sich Lernzuwächse je nach Untersuchungszeitraum unterscheiden (Christ et al., 2010), müssen die Datenerhebungen gezielt terminiert werden, wobei insbesondere Ferienzeiten und -effekte berücksichtigt werden sollten (Fink et al., 2015).

Normierung unter effektiver Förderung

Sofern Lernverlaufsnormen im Rahmen typischen Schulunterrichts erhoben werden, geben sie an, welche Lernzuwächse unter diesen Bedingungen zu erwarten sind. Eine weitere Norm wird allerdings notwendig, wenn ein Maßstab für gezielte Fördermaßnahmen bereitgestellt werden soll. Als einfache Heuristik geben Fuchs et al. (1993) an, dass eine durchschnittliche Entwicklung plus einer Standardabweichung zur Formulierung angemessener Ziele herangezogen werden kann. Um präziser abzuschätzen, welche Lernentwicklungen unter nachweislich effektiven Fördermethoden realistisch sind und damit eine anspruchsvolle Vergleichsnorm (auch Ziellinie) bereitzustellen, wäre eine Datenerhebung unter experimentellen Bedingungen (Experimental- vs. Kontrollgruppe) notwendig (Deno, Fuchs, Marston & Shin, 2001). Unter der Annahme, dass die Wirksamkeit einer Fördermethode von der Passung der Methode und den Lernvoraussetzungen abhängt, sollte eine „Fördernorm“ an einer homogenen Substichprobe erhoben werden und die Fördermethode genau beschrieben sein. Wie eine solche Substichprobe zusammengesetzt sein sollte, ist von dem jeweiligen Verfahren und dem Einsatzzweck abhängig. Lernverlaufsdiagnostik wird zwar vor allem im Response to Intervention-Ansatz für leistungsschwache Schülerinnen und Schüler verwendet (Huber & Grosche, 2012), grundsätzlich lässt sich mit

derartigen Verfahren aber auch evaluieren, ob Hochbegabte gemäß ihrem Potential optimal vom Unterricht profitieren (Hebbecke & Souvignier, 2016).

Analyseschritte bei der Normierung von Verfahren zur Lernverlaufsdiagnostik

Nachdem anhand eines sorgfältig gewählten Normierungsdesigns an einer repräsentativen Stichprobe Daten erhoben wurden, müssen Normwerte für die Lernverläufe bestimmt und Entscheidungsregeln dazu formuliert werden, ab wann eine Leistungsentwicklung so weit von der Norm abweicht, dass eine Anpassung der Instruktion notwendig ist. Dabei sind die folgend genannten Analyseschritte relevant.

Prüfung der Verlaufsform

Sofern ein Design mit mindestens vier Messzeitpunkten vorliegt, sollte in einem ersten Schritt anhand einer visuellen Inspektion und entsprechender Modellvergleiche (lineare oder quadratische Verläufe) überprüft werden, welche Verlaufsform die Daten ausreichend gut repräsentiert. Zwar dürften sich in den meisten Fällen lineare Verläufe finden, mitunter deuten quadratische Terme jedoch darauf hin, dass das Ausmaß der Lernzuwächse über die Zeit abnimmt (Fuchs et al., 1993; Strathmann & Klauer, 2010).

Berechnung der individuellen Lernzuwächse und Analyse auf Subtypen

Im nächsten Schritt müssen die individuellen Lernzuwächse anhand des passenden Modells bestimmt werden. Aufgrund der Abhängigkeit der Messungen bieten sich lineare oder generalisierte lineare Mischmodelle an (Fitzmaurice, Laird & Ware, 2011), bei denen individuelle Unterschiede in der Veränderung explizit und flexibel modelliert werden können.

¹ Diese Simulationstudien beziehen sich auf Tests, die nach der KTT konstruiert wurden. Möglicherweise gelten für anders konstruierte Tests (IRT) andere Zeiträume.

Wenn theoretische oder empirische Hinweise dazu vorliegen, dass sich Lernverläufe in Abhängigkeit von bestimmten Schülermerkmalen (z.B. Ausgangsniveau, Jahrgangsstufe, Geschlecht oder Muttersprache) unterscheiden, sollte dies explizit geprüft werden. Dazu lassen sich beispielsweise basierend auf den Testwerten zum ersten Testzeitpunkt Perzentilgruppen (1. bis 10. Perzentil, 11. bis 20. Perzentil, usw.) bilden, wobei Unterschiede in den Lernverläufen (z. B. mittels Quantilregression) zwischen den Gruppen explizit modellierbar sind (vgl. Silbergliitt & Hintze, 2007). Ein alternativer, exploratorischer Ansatz besteht darin, mit latenten Klassenanalysen unterschiedliche Verlaufstypen zu klassifizieren (Salaschek et al., 2014).

Bestimmung der durchschnittlichen Lernzuwächse und zugehöriger Vertrauensintervalle

In Abhängigkeit davon, wie unterschiedlich die gefundenen Lernzuwächse ausfallen, sollten durchschnittliche Lernzuwächse für entsprechende Subgruppen und die zugehörigen Vertrauensintervalle bestimmt werden. Bestenfalls werden dabei dynamische Vertrauensintervalle für jeden Messzeitpunkt angegeben, die berücksichtigen, dass Veränderungsmessungen mit zunehmender Dauer reliabler und valider werden (Christ et al., 2013). Je mehr Datenpunkte zur Verfügung stehen, desto genauer kann der Lernzuwachs (slope) bestimmt werden und desto kleiner wird das Vertrauensintervall für den slope. Insofern sollte für die jeweilige Anzahl an Messungen ein jeweils spezifisches Vertrauensintervall für den Lernzuwachs angegeben werden.

Entwicklung einer Entscheidungsregel

Eine wesentliche Frage ist, ab wann eine verlässliche Aussage über die Lernentwicklung möglich ist, so dass eine pädagogische Entscheidung getroffen werden kann und

was eine sinnvolle Interpretation der bis dahin erhobenen Informationen ist. Generell werden dazu die beobachteten Lernfortschritte mit dem zu erwartenden Lernfortschritt (sogenannte Ziellinie) verglichen (Deno et al., 2001) und es gibt Entscheidungsregeln hinsichtlich a) des Beibehaltens der Instruktion (maintain instruction), b) der Instruktionsanpassung (change instruction) und c) der Erhöhung der Ziellinie (increase goal). Die Entscheidungsregeln lassen sich grundsätzlich den beiden Kategorien der Datenpunkt- und Trendregeln zuordnen (Ardoin, Christ, Morena, Cormier & Klingbeil, 2013). Während bei der Datenpunkt-Regel typischerweise die letzten drei Messungen mit der erwarteten Entwicklung verglichen werden, wird bei Trendregeln der tatsächliche Lernzuwachs (ermittelt z.B. über regressionsanalytische Verfahren) mit der erwarteten Entwicklung verglichen. Grundsätzlich scheinen Entscheidungsregeln, die auf dem Lernzuwachs basieren, zu gültigeren Entscheidungen zu kommen, als Regeln, bei denen lediglich die letzten drei Messzeitpunkte mit der Norm verglichen werden (Van Norman & Christ, 2016).

Insgesamt ist eine verallgemeinerte Empfehlung für Entscheidungsregeln schwierig. Das resultiert nicht zuletzt daraus, dass zwar einerseits das Vertrauensintervall des Lernfortschritts mit zunehmenden Messzeitpunkten kleiner wird, die Güte einer Entscheidung gleichzeitig jedoch auch davon abhängt, wie sehr die Lernentwicklung von der Norm abweicht (Van Norman & Christ, 2016). Je größer der Unterschied zwischen der individuellen Lernentwicklung und der Norm ist, desto früher kann diese Abweichung auch bei noch großen Vertrauensintervallen des slope identifiziert werden. Nach Christ et al. (2013) muss der Lernzuwachs wenigstens doppelt so groß sein wie der Standardmessfehler des Zuwachses², was in ihren Simulationsstudien erst nach et-

² Da das 1,96-fache des Standardmessfehlers das Konfidenzintervall des Zuwachses beschreibt, ist insofern von einer reliablen Veränderung auszugehen.

wa acht bis zehn Wochen Verlaufsdiagnostik der Fall war. Bis verlässliche Schätzungen des Lernzuwachses möglich sind, sollten Lernverläufe entsprechend vorsichtig interpretiert werden. In jedem Fall sollten explizite Regeln zur Interpretation angegeben werden, da die Bereitstellung expliziter Entscheidungsregeln als wesentlich für die Effektivität der Lernverlaufsdiagnostik betrachtet wird (Stecker, Fuchs & Fuchs, 2005) und Lehrkräfte Schwierigkeiten bei der freien Interpretation entsprechender Daten haben (Zeuch, Förster & Souvignier, 2017).

Zusammenfassung

Der wesentliche Unterschied zur Statusdiagnostik und gleichzeitig die zentrale Anforderung an die Normierung von Verfahren zur Lernverlaufsdiagnostik besteht darin, dass Normen für die Veränderung bereitgestellt werden müssen. Zudem ist zu beachten, dass die Anwendungssituationen für Lernverlaufsdiagnostik sehr vielfältig sind. Damit Normdaten im Kontext der Lernverlaufsdiagnostik als geeigneter Vergleichsmaßstab oder für eine sinnvolle Zielsetzung genutzt werden können, müssen sie in enger Passung zum jeweiligen Anwendungszweck erhoben worden sein. Dies erfordert eine besondere Sorgfalt bei der Auswahl repräsentativer Stichproben und unterrichtlicher Settings sowie eine präzise Dokumentation der Umstände, unter denen die Lernverläufe der Normierungsstichprobe erfasst wurden.

Der Beitrag beschreibt darüber hinaus zwei weitere entscheidende Herausforderungen, die mit einer derartigen Normierung einhergehen. So zeigen empirische Befunde, dass die Muster der Lernverläufe unterschiedlich ausfallen können. Zum Beispiel finden sich in Abhängigkeit von der erfassten Kompetenz, von den verwendeten Verfahren sowie weiteren Schülermerkmalen unterschiedliche Lernverlaufsmuster, die neben linearen Trends auch andere Verläufe aufweisen. Zum anderen sind bei unterschiedlichen Kindern (beispielsweise in

Abhängigkeit vom Leistungsniveau) verschieden hohe Lernzuwächse zu erwarten. Entsprechend sollten Art und Anzahl unterschiedlicher Verläufe bei der Normierung explizit untersucht werden. Darüber hinaus erfordert der formative Charakter der Lernverlaufsdiagnostik explizite Entscheidungsregeln für das Beibehalten oder Ändern unterrichtlicher Angebote für die Anwender, wobei die Reliabilität und Validität der Veränderungsmaßnahmen über die Bestimmung dynamischer Vertrauensintervalle berücksichtigt werden sollten.

Literatur

- Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C. & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding Curriculum-based Measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology, 51*, 1-18.
- Asendorpf, J. B. (2000). Idiographische und nomothetische Ansätze in der Psychologie. *Zeitschrift für Psychologie, 208*, 72-90.
- Christ, T. J., Monaghan, B. D., Zopluoglu, C. & Van Norman, E. R. (2012). Curriculum-based measurement of oral reading: Evaluation of growth estimates derived with pre-post assessment methods. *Assessment for Effective Intervention, 38*, 139-153.
- Christ, T. J., Silbergliitt, B., Yeo, S. & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review, 39*, 447-462.
- Christ, T. J., Zopluoglu, C., Monaghan, B. D. & Van Norman, E. R. (2013). Curriculum-based Measurement of Oral Reading: Multi-study evaluation of schedule, duration, and dataset quality on progress monitoring outcomes. *Journal of School Psychology, 51*, 19-57.
- Deno, S. L., Fuchs, L. S., Marston, D. & Shin, J. (2001). Using curriculum-based measu-

- ment to establish growth standards for students with learning disabilities. *School Psychology Review*, 30, 507-524.
- Fink, A., Luttenberger, S., Kramme, A., Macher, D., Papousek, I., Weiss, E. & Paechter, M. (2015). Die Veränderung kognitiver Fähigkeiten über die Sommerferien. *Psychologie in Erziehung und Unterricht*, 62, 303-315.
- Fitzmaurice, G. M., Laird, N. M. & Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed.). Hoboken, NJ: Wiley.
- Förster, N. & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, 32, 91-100.
- Förster, N. & Souvignier, E. (2015). Effects of providing teachers with information about their students reading progress. *School Psychology Review*, 44, 60-75.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L. & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, 22, 27-48.
- Grosche, M. (2011). Effekte einer direkt-instruktiven Förderung der Lesegenauigkeit. *Empirische Sonderpädagogik*, 3, 147-161.
- Hebbecker, K. & Souvignier, E. (2016). Lernverlaufsdagnostik zur Unterstützung für individuelle (Begabungs-)Förderung: internetbasierte Lernverlaufsdagnostik mit dem System quop. *Journal für Begabtenförderung*, 16, 29-38.
- Huber, C. & Grosche, M. (2012). Das response-to-intervention-Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. *Zeitschrift für Heilpädagogik*, 63, 312-322.
- Keller-Margulis, M. A., Clemens, N. H., Im, M. H., Kwok, O. M. & Booth, C. (2012). Curriculum-based measurement yearly growth rates: An examination of English Language Learners and native English speakers. *Learning and Individual Differences*, 22, 799-805.
- Salaschek, M., Zeuch, N. & Souvignier, E. (2014). Mathematics growth trajectories in first grade: Cumulative vs. compensatory patterns and the role of number sense. *Learning and Individual Differences*, 35, 103-112.
- Silbergitt, B. & Hintze, J. M. (2007). How much growth can we expect? A conditional analysis of R—CBM growth rates by level of performance. *Exceptional Children*, 74, 71-84.
- Stecker, P. M., Fuchs, L. S. & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools*, 42, 795-819.
- Strathmann, A. M. & Klauer, K. J. (2010). Lernverlaufsdagnostik: Ein Ansatz zur längerfristigen Lernfortschrittsmessung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42, 111-122.
- Van Norman, E. R. & Christ, T. J. (2016). Curriculum-based measurement of reading: Accuracy of recommendations from three-point decision rules. *School Psychology Review*, 45, 296-309.
- Voß, S., Blumenthal, Y., Diehl, K., Mahlau, K., Sikora, S. & Hartke, B. (2014). Rügener Inklusionsmodell (RIM) – Effekte eines Beschulungsansatzes nach dem Response to Intervention-Ansatz auf die Rechen- und Leseleistungen von Grundschulkindern. *Empirische Sonderpädagogik*, 6, 114-132.
- Zeuch, N., Förster, N. & Souvignier, E. (2017). Assessing teachers' competencies to read and interpret graphs from learning progress assessment: Results from tests and interviews. *Learning Disabilities Research & Practice*, 32, 61-70.

Natalie Förster

Westfälische Wilhelms-Universität
Münster
Fliednerstraße 21
48149 Münster
natalie.foerster@uni-muenster.de

Erstmalig eingereicht: 01.03.2017

Überarbeitung eingereicht: 13.06.2017

Angenommen: 21.07.2017