

**Erkenntnisse aus Large-Scale Assessments als empirischer  
Bezugspunkt für bildungspolitische Entscheidungen**

Habilitation  
an der TUM School of Education

Dr. Christine Sälzer

November 2015

# Inhalt

|  |    |
|--|----|
| Inhalt.....  | 2  |
| Erkenntnisse aus Large-Scale Assessments als empirischer Bezugspunkt ..... |    |
| für bildungspolitische Entscheidungen.....                                 | 3  |
| 1. Kontext und Relevanz der Thematik.....                                  | 3  |
| 2. Zusammenfassung der Studien und ihres Erkenntnisgewinns.....            | 6  |
| 3. Eigenleistungen im Kontext der Fachpublikationen .....                  | 21 |
| Literaturverzeichnis.....  | 23 |
| Lebenslauf .....   | 30 |
| Publikationen und Vorträge.....  | 34 |
| Fachpublikationen der Habilitation .....                                   | 40 |

# **Erkenntnisse aus Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen**

## **1. Kontext und Relevanz der Thematik**

Befunde aus Large-Scale Assessments wie etwa der PISA-Studie liefern empirische Hinweise auf Stärken und Schwächen von Bildungssystemen. Bereits seit dem Konstanzer Beschluss (KMK, 1997) gehört die regelmäßige Teilnahme an international vergleichenden Schulleistungsstudien zur KMK-Gesamtstrategie für ein Bildungsmonitoring (KMK, 2006; KMK, 2015). Vor der Beteiligung an der *Third International Mathematics and Science Study* (TIMSS) im Jahr 1995 hatte Deutschland mehrere Jahrzehnte lang an keiner international vergleichenden Schulleistungsstudie teilgenommen, weshalb die mittelmäßigen Ergebnisse aus TIMSS 1995 und wenige Jahre später aus PISA 2000 die heute als „PISA-Schock“ (Roeder, 2003) bekannte Reaktion auslösten. Die gravierenden Probleme, die anhand der Befunde aus TIMSS und PISA sichtbar wurden, waren durch den langen Verzicht auf einen internationalen Vergleich (Prenzel, 2005) und das empirisch nicht hinterfragte Vertrauen in die Leistungsfähigkeit des deutschen Schulsystems (Gräsel, 2011) mit erheblicher Verzögerung entdeckt worden. Aus wissenschaftlicher Sicht markieren der Konstanzer Beschluss und die damit verbundene kontinuierliche Beteiligung an Large-Scale Assessments also einen Wendepunkt in Deutschland, mit dem eine Entwicklung hin zu einer an Daten und Evidenz interessierten und orientierten Politik auch im Bildungssektor begann (Bromme et al., 2014; Ditton, 2000; Tillmann et al., 2008). Eine Bildungspolitik, die bei Entscheidungen und Steuerungsprozessen empirische Befundlagen berücksichtigt, greift auf eine bestimmte, evidenzbasierte Form des Wissens über das Bildungssystem zurück (Kuper, 2006). Seit der Beteiligung Deutschlands an internationalen Large-Scale Assessments wird dieses Wissen quantitativ, standardisiert und vergleichend gewonnen.

Internationale Schulleistungsstudien dienen dabei vor allem zwei Funktionen: Monitoring und Benchmarking (Seidel & Prenzel, 2008). Beides impliziert Vergleiche mit bestimmten Standards (Monitoring) beziehungsweise mit der Struktur, den Prozessen und Ergebnissen anderer Bildungssysteme (Benchmarking). Zentrale Erkenntnisse aus diesen Studien betreffen zum einen das Kompetenzniveau der teilnehmenden Schülerinnen und Schüler und dessen Verteilung sowie Rahmenbedingungen, die Einfluss auf die Kompetenzentwicklung ausüben können. Zum anderen liefert der internationale Vergleich der Leistungsfähigkeit von Schülerinnen und Schülern zu bestimmten Zeitpunkten in der Schullaufbahn weitere Vergleichsperspektiven, die Trends im Sinne von Entwicklungen über die Zeit erfassen und abbilden können (Rutkowski & Prusinski, 2011). So können neben der internationalen Verankerung von Schülerleistungen Vergleiche dieser Leistungen über mehrere Erhebungsrunden angestellt, mögliche Wirkungen von Maßnahmen abgeschätzt und problematische Entwicklungen frühzeitig erkannt werden (Drechsel & Prenzel, 2008; Drechsel et al., 2014). Die mit dem Konstanzer Beschluss eingeleitete „empirische Wende“ (Buchhaas-Birkholz, 2009;

Lange, 2008) der deutschen Bildungspolitik betrifft demnach zwei Aspekte: erstens die langfristige theoriegeleitete und systematische empirische Messung von Kompetenzen und zweitens die Rationalität des bildungspolitischen Handelns in Deutschland. Gemeinsam betrachtet bedeuten diese beiden Aspekte, dass empirisch gewonnene Daten und deren Interpretation im politischen Diskurs als Bezugspunkt für die Begründung von Entscheidungen verwendet werden können (Bieber et al., 2014; Bromme et al., 2014).

Large-Scale Assessments und insbesondere PISA werden in der vorliegenden Habilitationsschrift als ein Instrument des Bildungsmonitorings aufgefasst, dessen Ergebnisse überwiegend beschreibendes Wissen darstellen (vgl. etwa Prenzel, 2009). Dieses beschreibende Wissen ist ein Bezugspunkt für Schlussfolgerungen und anknüpfende Forschung, so dass neben der Erfassung von Problemen im Bildungsbereich auch mögliche Ansatzpunkte für Veränderungsprozesse identifiziert werden können (Prenzel, 2005). Befunde aus Large-Scale Assessments sowie aus daran anschließender Forschung sind also Gegenstand der Interaktion von Bildungsforschung und Bildungspolitik. In diesem Kontext verknüpft die vorliegende publikationsbasierte Habilitation zwei Aspekte: die Bereitstellung differenzierter Analysen des Bildungsstands von Schülerinnen und Schülern in Deutschland und deren Entstehens- und Entwicklungsbedingungen sowie die wissenschaftliche Erweiterung von Daten aus Large-Scale Assessments für die Bearbeitung von spezifischen Fragestellungen. Diese Fragestellungen tragen dazu bei, Erkenntnisse über Teilpopulationen von Schülerinnen und Schülern zu gewinnen, die im Rahmen der PISA-Studie üblicherweise nicht repräsentativ abgebildet werden können.

Zusammengefasst lassen sich Large-Scale Assessments in grundlegende Funktionen der empirischen Bildungsforschung als Forschungsfeld einordnen, in dem Voraussetzungen, Prozesse und Ergebnisse von Bildung sowie deren Beziehungen erfasst werden (Gräsel, 2011; Lüders, 2004; Prenzel, 2005). Die beschreibende Funktion von Large-Scale Assessments, deren Hauptziel die Identifikation von Problemen ist, wird im Kontext der empirischen Bildungsforschung als regelmäßiges Monitoring verwendet und über gezielte Zusammenhangsanalysen auch in eine prädiktive Funktion überführt. Zu diesem Zweck werden in der empirischen Bildungsforschung mehrere methodische Ansätze genutzt, wie die Erhebungen in Large-Scale Assessments für spezifische Fragestellungen erweitert werden können. Drei zentrale Vorgehensweisen, die im Rahmen von PISA in Deutschland bereits mehrfach gewählt wurden, sind die Ziehung von Ergänzungsstichproben (*Oversamples*), die Erweiterung der eingesetzten Befragungs- und Testinstrumente (vgl. etwa Prenzel et al., 2008) und die längsschnittliche Erweiterung durch Messwiederholungen nach einem Jahr (vgl. etwa Baumert et al., 2002; Prenzel et al., 2006; Prenzel & Baumert, 2008).

Daran anknüpfend wird die übergeordnete Thematik der Habilitation in zwei Forschungsbereiche gegliedert, die in Bezug auf die Nutzung von Large-Scale Assessments unterschiedliche Schwerpunkte setzen: Erstens die Rolle von Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen und zweitens differenzierende Befunde aus Large-Scale

Assessments für Teilpopulationen von Schülerinnen und Schülern, die ohne eine Erweiterung der Stichprobe nicht repräsentativ abgebildet werden könnten. Der erste Forschungsbereich konzentriert sich auf eine der Hauptaufgaben von Large-Scale Assessments: Monitoring und Benchmarking. Die vertiefenden Analysen und nationalen Berichte zu PISA 2012 sowie die wissenschaftliche Reflektion der Rezeption früherer Befunde aus Large-Scale Assessments in Deutschland bilden den Schwerpunkt dieses Bereichs. Der zweite Forschungsbereich greift auf Daten aus einer Ergänzungsstudie zu PISA 2012 zurück, die sich gezielt einer Teilpopulation der in PISA untersuchten Kohorte fünfzehnjähriger Schülerinnen und Schüler widmete: Jugendlichen mit sonderpädagogischem Förderbedarf.

Diese Gruppe von Schülerinnen und Schülern ist unter anderem deshalb von besonderem bildungspolitischem Interesse, weil derzeit noch etwa jede zehnte allgemeinbildende Schule in Deutschland eine Förderschule ist (Autorengruppe Bildungsberichterstattung, 2014). Das Sonderschulsystem in Deutschland war zunächst vorwiegend in den 1970er Jahren ausgebaut worden, um der erst damals eingeführten Schulpflicht für Kinder mit einer Behinderung nachzukommen (Gebhardt et al., 2013). Mit der Ratifizierung der UN-Behindertenrechtskonvention in Deutschland im Jahr 2009 werden Schülerinnen und Schüler mit diagnostiziertem sonderpädagogischem Förderbedarf nun zunehmend integrativ beziehungsweise inklusiv an Regelschulen beschult (Autorengruppe Bildungsberichterstattung, 2014; Gebhardt et al., 2014). In Bezug auf die Umsetzung der Inklusion in Deutschland wird derzeit intensiv diskutiert. Die Debatte um Inklusion benötigt jedoch fundiertes Wissen über Zusammenhänge unterschiedlicher Schulformen mit der Leistungsentwicklung und Lernbedingungen bei Schülerinnen und Schülern mit besonderem Förderbedarf. Hinter dem Vergleich von Schülerinnen und Schülern mit besonderem Förderbedarf an Förderschulen und solchen an allgemeinbildenden Sekundarschulen steht unter anderem die Frage nach unterschiedlichen Bedingungen und Konsequenzen in den Beschulungsorten Regelschule und Förderschule, der bisher in erster Linie für die Grundschulstufe bearbeitet wurde (Kocaj et al., 2014; Kocaj et al., 2015). Gleichzeitig ist die empirische Evidenz über einzelne Erfolgs- und Misserfolgskriterien oder mögliche Auswirkungen noch immer relativ ungesichert (Blanck et al., 2013; Pfahl, 2012).

Noch offen sind derzeit Befunde zum Kompetenzniveau und der Kompetenzverteilung von Sekundarschülerinnen und -schülern mit sonderpädagogischem Förderbedarf, insbesondere hinsichtlich Kompetenzen wie Mathematik, Lesen und Naturwissenschaften, die im Sinne des *Literacy*-Konzepts in PISA definiert wurden (vgl. etwa OECD, 2013a). Das *Literacy*-Konzept stellt mit der Teilhabe und der Anschlussfähigkeit dieser Kompetenzen zwei Aspekte funktionaler Grundbildung in den Mittelpunkt (Sälzer & Prenzel, 2013), die gerade in Bezug auf die Debatte um Inklusion einen empirischen Bezugspunkt benötigen. In Deutschland steht dabei auch die bundesweit repräsentative Beschreibung von Leistungsprofilen von an Förderschulen und an den übrigen allgemeinbildenden Sekundarschulen unterrichteten Schülerinnen und Schülern noch weitgehend aus (vgl. etwa Wocken & Gröhlich, 2009). Dies bringt auch methodische Herausforderungen mit sich. Gerade beim internationalen Vergleich von

Schülerinnen und Schülern mit besonderem Förderbedarf zeigt sich in der einschlägigen Literatur wiederholt, dass die Definition und damit auch die Identifikation und letztlich die Diagnostik für die Zuschreibung eines besonderen Förderbedarfs hochgradig zwischen Staaten variiert (vgl. exemplarisch Demeris et al., 2007; Myklebust, 2002; Petermann & Petermann, 2006). Im föderalen Bildungssystem Deutschlands liegt zwar eine rechtliche Grundlage seitens der Kultusministerkonferenz vor, diese wird allerdings bundeslandspezifisch sehr unterschiedlich ausgestaltet (KMK, 1994). Dadurch ist die Klassifikation und Erfassung eines sonderpädagogischen Förderbedarfs bei Schülerinnen und Schülern nicht einheitlich und damit letztlich auch nicht eindeutig. Folglich stellt die Berücksichtigung von Schülerinnen und Schülern mit besonderem Förderbedarf nicht nur eine methodische Herausforderung bei internationalen Vergleichen dar, sondern auch bei Studien, die innerhalb Deutschlands mehrere oder alle Bundesländer umfassen. Um sowohl innerhalb Deutschlands als auch im Kontext international vergleichender Schulleistungsstudien wie PISA eine belastbare Grundlage für die Definition und Erfassung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf zu schaffen, widmete sich ein inhaltlicher Schwerpunkt der Arbeiten im zweiten Forschungsbereich einer Bestandsaufnahme von Akteuren und Prozessen der Feststellung und Zuweisung eines sonderpädagogischen Förderbedarfs in den 16 deutschen Bundesländern. Diese Beschreibung der Diagnostik eines sonderpädagogischen Förderbedarfs liegt außerhalb der Möglichkeiten und Interessen von PISA selbst. Ergänzt wurde dieser Aspekt durch zwei Aufsätze zur empirischen Analyse der Kompetenzen von Schülerinnen und Schülern an Förderschulen in Deutschland im Rahmen einer Stichprobenerweiterung zu PISA 2012 an 49 Förderschulen.

## **2. Zusammenfassung der Studien und ihres Erkenntnisgewinns**

Die Habilitation umfasst eine Reihe von Fachpublikationen<sup>1</sup>, welche in Tabelle 1 aufgeführt sind. Die ersten sieben Beiträge sind dem ersten Forschungsbereich zugeordnet, der die Monitoringfunktion von PISA in den Mittelpunkt rückt und die daraus gewonnenen Daten als empirischen Bezugspunkt für bildungspolitische Entscheidungen aufgreift. Drei weitere Aufsätze sind dem zweiten Forschungsbereich zugeordnet und behandeln empirische Fragestellungen zu den Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Deutschland. Insofern baut der zweite Forschungsbereich auf dem ersten auf und entwickelt vertiefende Fragestellungen zur Deskription von Merkmalen der PISA-Schülerpopulation und einer Teilpopulation.

In den folgenden Abschnitten werden die beiden Forschungsbereiche weiter ausgeführt und die Ergebnisse sowie Erkenntnisse der jeweils zugeordneten Aufsätze verdichtet. Anschließend wird jeweils die eigene Leistung im Kontext der Fachpublikationen dargelegt.

---

<sup>1</sup> Gemäß § 9 (2) 2 der Habilitationsordnung der Technischen Universität München vom 9. Dezember 2003, in der Fassung der Änderungssatzung vom 13. Dezember 2005.

Tabelle 1. Übersicht der Fachpublikationen im Rahmen der Habilitation

---

### **Forschungsbereich I**

#### **Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen**

1. Sälzer, C. & Prenzel, M. (2013). PISA 2012 – eine Einführung in die aktuelle Studie. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 11-46).Münster: Waxmann.
2. Heine, J.-H., Sälzer, C., Borchert, L., Sibberns, H. & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 309-346). Münster: Waxmann,.
3. Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenz im internationalen Vergleich. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 47-98). Münster: Waxmann.
4. Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 217-244). Münster: Waxmann.
5. Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 155-188). Münster: Waxmann.
6. Sälzer, C. & Prenzel, M. (2014). Looking Back at Five Rounds of PISA: Impacts on Teaching and Learning Mathematics in Germany. *Solsko Polje (The School Field)*, 25 (5-6), 53-72.
7. Sälzer, C., Schiepe-Tiska, A. & Prenzel, M. (revise & resubmit). Internationale Schulleistungsstudien. In: M. Gläser-Zikuda, M. Harring, C. Rohlf's (Hrsg.), *Handbuch Schulpädagogik*. Münster: Waxmann.

---

### **Forschungsbereich II**

#### **Befunde aus Large-Scale Assessments für bestimmte Teilpopulationen von Schülern**

8. Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In: P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Forschungsstand, Herausforderungen und Perspektiven* (S. 129-152).Wiesbaden: Springer.
  9. Gebhardt, M., Sälzer, C., Mang, J., Müller, K., & Prenzel, M. (2015). Performance of Students with Special Educational Needs in Germany. Findings from PISA 2012. *Journal of Cognitive Education and Psychology*, 14 (3), 343-356.
  10. Müller, K., Prenzel, M., Sälzer, C., Mang, J., Heine, J.-H. & Gebhardt, M. (revise & resubmit). Wie schneiden Schülerinnen und Schüler an Sonder- und Förderschulen bei PISA ab? Analysen aus der PISA 2012-Zusatzerhebung zu Jugendlichen mit sonderpädagogischem Förderbedarf. Eingereicht bei *Unterrichtswissenschaft*.
-

### ***Forschungsbereich 1:***

#### ***Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen***

Internationale Schulleistungsuntersuchungen werden meist unter den Zielsetzungen Monitoring und Benchmarking durchgeführt (Seidel & Prenzel, 2008). Sowohl Monitoring als auch Benchmarking bedeutet, dass Ergebnisse mit bestimmten Standards (*Monitoring*) beziehungsweise mit anderen Bildungssystemen (*Benchmarking*) verglichen werden. Im Sinne des Monitorings liefert PISA zunächst Zustandsbeschreibungen und darüber hinaus eine Datengrundlage für Überlegungen zur Veränderung und Verbesserung von Bildungsprozessen, während die Benchmarkingfunktion relative Stärken und Schwächen im Vergleich zu anderen Bildungssystem sichtbar macht (Drechsel et al., 2009). Der Benchmarking-Aspekt betont so gesehen also den Vergleich mit anderen Staaten und Bildungssystemen, während der Monitoring-Aspekt stärker den Einzelstaat und dabei überwiegend die Entwicklung über die Zeit (Trend) im Blick hat. Schülerkompetenzen werden in Large-Scale Assessments in der Regel modellbasiert anhand eines umfangreichen Instrumentariums gemessen. PISA folgt dabei dem Prinzip des Multi-Matrix-Designs (vgl. etwa Heine et al., 2013), das basierend auf Annahmen der *Item-Response-Theory* (Reise & Revicki, 2014; Yamamoto & Mazzeo, 1992) die Antworten von Schülerinnen und Schülern auf die Testfragen und Fragebögen als Grundlage für die Schätzung durchschnittlicher Kompetenzen auf Aggregatebene heranzieht (Sälzer, 2015). Das Design und die Testkonzeption von PISA sehen demnach also keine individuelle Leistungsdiagnostik vor, sondern Aussagen über durchschnittliche Ausprägungen und Streuungen ausgewählter Maße wie Kompetenzen oder Einstellungen (vgl. etwa Carstensen et al., 2004).

Die für die Habilitation vorgelegten Aufsätze erweitern die Berichterstattung der OECD zu PISA 2012 in zweierlei Hinsicht: (1) in einer *vertiefenden Beschreibung und Analyse* von speziell für die Bildungspolitik und –praxis Deutschlands relevanten Aspekten des Bildungssystems und (2) in der Vertiefung des Kenntnisstandes im Bildungsmonitoring auf der Sekundarstufe I durch die wissenschaftliche Einordnung von *Entwicklungen* in Deutschland über mehrere PISA-Erhebungsrunden hinweg. In Forschungsbereich I sind Fragen zum Kompetenzniveau, seiner Verteilung und Entwicklungsbedingungen der in PISA untersuchten Population fünfzehnjähriger Schülerinnen und Schüler in Deutschland zentral. Mit Blick auf die Untersuchung von Trends stehen darüber hinaus Fragen nach dem Umgang mit in früheren PISA-Erhebungsrunden identifizierten Problemlagen wie Disparitäten innerhalb Deutschlands im Mittelpunkt des Erkenntnisinteresses.

#### **Vertiefende Beschreibung relevanter Indikatoren für die deutsche Bildungspolitik und -praxis.**

Als relativ junges Forschungsfeld ist die empirische Bildungsforschung aktuell geprägt durch eine große Vielfalt an disziplinären, theoretischen und methodischen Zugängen (Gräsel, 2011; Prenzel, 2005). Large-Scale Assessments wie PISA sind ein quantitativer Ansatz zur Gewinnung von Erkenntnissen über die Rahmenbedingungen und Ergebnisse vorwiegend institutionalisierter Bildungsprozesse, die sich auf sogenannte Domänen als inhaltliche Lern- und Handlungsbereiche beziehen (Klieme &

Leutner, 2006). Als Bezugspunkt für evidenzbasierte politische Steuerungsprozesse (Bromme et al., 2014; Slavin, 2002) verbindet die nationale Berichterstattung zu PISA die vertiefende Beschreibung relevanter Indikatoren für Bildungsergebnisse (Monitoring) mit einem internationalen Vergleich (Benchmarking). **Beitrag 1** (Sälzer & Prenzel, 2013) ist als Einleitung des nationalen Berichtsbandes zu PISA 2012 in diesem Kontext verankert, indem die Zielsetzung der PISA-Studie sowie die Erhebung und Auswertung der berichteten Daten gezielt mit Blick auf Deutschland systematisch dargestellt werden. Im Mittelpunkt des Beitrags steht neben der Frage nach der Entstehung der PISA-Daten auch diejenige nach der Absicherung der internationalen Vergleichbarkeit der Ergebnisse. Für die fundierte Auswertung der Daten und die Interpretation der Befunde sind der Entstehungskontext sowie die beteiligte Stichprobe, das Design, die Testkonzeption und die adäquate Übersetzung der Testinstrumente ebenso zentral wie die Rahmenbedingungen für die Durchführung von PISA in Deutschland, etwa in Bezug auf die Verpflichtung zur Bearbeitung der Kontextfragebögen. Zwar werden in Deutschland wie in den übrigen PISA-Teilnehmerstaaten die international vorgegebenen Prozeduren und Standards verbindlich eingehalten, jedoch existieren darüber hinaus einige Besonderheiten. So muss bei der Stratifizierung der Untersuchungspopulation berücksichtigt werden, dass die 16 Länder entsprechend dem Anteil an fünfzehnjährigen Schülerinnen und Schülern repräsentiert sind, wobei diese jeweils unterschiedliche Schulformen zur Auswahl haben. Außerdem ist lediglich die Teilnahme am Kompetenztest durch die KMK bundesweit verbindlich geregelt. In Bezug auf die Kontextfragebögen gelten die jeweiligen Bestimmungen der Länder. Insofern geht der erste Beitrag deutlich über die internationale Berichterstattung der OECD hinaus (vgl. u. a. OECD, 2013b; OECD, 2014a), deren Perspektive im Vergleich zu einer nationalen Berichterstattung wesentlich stärker die Ebene der staatenübergreifenden Beschreibung der Studie fokussiert. In dieser Berichterstattung werden oftmals bivariate Zusammenhänge oder Rangreihen von Staaten aufgrund einzelner Items aus Fragebogenskalen präsentiert. Der Verzicht auf eine differenzierte Beschreibung der jeweiligen organisatorischen und institutionellen Gegebenheiten in den teilnehmenden Staaten kann jedoch Fehlinterpretationen begünstigen. Insofern ist die systematische Beschreibung der Entstehung und Verarbeitung der Daten in PISA 2012 essentiell für eine belastbare Verwendung der Ergebnisse als Bezugspunkt für bildungspolitische Entscheidungen in Deutschland.

Large-Scale Assessments sind ein quantitativer Ansatz zur Gewinnung von Erkenntnissen über die Rahmenbedingungen und Ergebnisse überwiegend institutionalisierter Bildungsprozesse. Sie stellen öffentlich verfügbare Datensätze bereit, die sowohl als Bezugspunkt für evidenzbasierte politische Steuerungsprozesse (Bromme et al., 2014; Slavin, 2002) als auch für empirische Forschung verwendet werden (Rammstedt & Spinath, 2013). **Beitrag 2** (Heine et al. 2013) unterstützt dies, indem die technischen und methodischen Grundlagen für die Vorbereitung, Erhebung und Verwertung der in PISA erzeugten Daten transparent erläutert und vor dem Hintergrund aktueller methodischer Erkenntnisse erörtert werden. In Ergänzung zu Beitrag 1 liefert Beitrag 2 eine Beschreibung des methodischen Zugangs in PISA sowie des eingesetzten Instrumentariums, die hinreichend exakt ist und auch für Nicht-

Experten verständlich bleibt. Um die durch PISA gewonnenen Daten in ihrer Tiefe nutzen und korrekte Schlussfolgerungen ziehen zu können, ist eine fundierte Kenntnis der Struktur und Zusammensetzung des Datensatzes ebenso unerlässlich wie das Wissen über die Rahmenbedingungen und Prozeduren der Datenerhebung. Beitrag 2 knüpft also an Beitrag 1 an und führt zunächst in die Ziehung der Schul- und Schülerstichprobe ein. In diesem Zusammenhang geht der Beitrag darauf ein, inwieweit ein Gewichtungungsverfahren neben unterschiedlichen Ziehungswahrscheinlichkeiten von Schulen und Schülern auch Ausfälle auf Schul- und Schülerebene korrigiert (vgl. u. a. Fay, 1989). Im nächsten Schritt wird erläutert, wie die Kompetenzen in PISA gemessen und letztlich geschätzt werden, wobei verschiedene Methoden zur Parameterschätzung verglichen und diskutiert werden. Als das am besten zum Testdesign von PISA 2012 passende Verfahren wurde die *Marginal-Maximum-Likelihood*-Methode eingesetzt (Bock & Lieberman, 1970; Bock & Aitkin, 1981). Aufbauend auf der Argumentation zur Wahl des Schätzverfahrens für die Item- und Personparameter für die Kompetenzmessung wird abschließend auf das Verfahren zur Konstruktion von Kompetenzstufen eingegangen sowie auf die Abbildung von Trends über die bislang abgeschlossenen fünf PISA-Erhebungsrunden. Der Beitrag ist eine wichtige Ergänzung zum sogenannten *Technical Report* der OECD zu PISA 2012 (OECD, 2014b), da er zum einen konkrete Informationen zur Stichprobe in Deutschland gibt und zum anderen diejenigen methodischen Aspekte auswählt und diskutiert, die der nationalen Berichterstattung zu PISA in Deutschland zu Grunde liegen und neben Gemeinsamkeiten zu früheren PISA-Erhebungsrunden auch Änderungen thematisieren.

Ein zentraler Aspekt des Bildungsmonitorings ist die wiederholte Erfassung von Indikatoren über einen längeren Zeitraum und damit verbunden die Beobachtung von Trends. Daran knüpfen die **Beiträge 3 und 4** an, indem die durchschnittlichen Kompetenzen fünfzehnjähriger Schülerinnen und Schüler in Deutschland in den Domänen Mathematik und Deutsch untersucht werden. **Beitrag 3** (Sälzer et al., 2013) beschreibt das theoretische Konstrukt und die empirischen Ergebnisse zur mathematischen Kompetenz, die in PISA 2012 nach 2003 zum zweiten Mal den inhaltlichen Schwerpunkt bildete. Im Mittelpunkt steht neben der detaillierten Definition der im PISA-Mathematiktest gemessenen Dimensionen mathematischer Kompetenz im Sinne des *Literacy*-Konzepts und Fragen der Messqualität und Vergleichbarkeit (OECD, 2013a) auch die Frage nach dem durchschnittlichen Kompetenzniveau der Schülerinnen und Schüler in Deutschland sowie dessen Entwicklung seit PISA 2003. Ein wichtiger deskriptiver Befund ist, dass die Fünfzehnjährigen in Deutschland in PISA 2012 inzwischen Leistungen erzielen, die signifikant über dem OECD-Durchschnitt liegen. In Erweiterung zu den in der theoretischen Rahmenkonzeption der OECD (OECD, 2013a) erläuterten Teilkomponenten der Domäne Mathematik wird in Beitrag 3 ein konzeptueller Bezug zu den deutschen Bildungsstandards in Mathematik für den Mittleren Schulabschluss (KMK, 2003) hergestellt. Die Testkonzeption in PISA basiert auf Anschlussfähigkeit und Teilhabe und strebt nicht primär curriculare Validität an (vgl. etwa Baumert et al., 2001; Klieme et al., 2007). Gleichwohl gibt es inhaltliche Überschneidungen des Curriculums mit dem im *Literacy*-Konzept definierten Kompetenzbegriff (vgl. etwa Ehmke et al., 2006).

Die Herausarbeitung der Gemeinsamkeiten und Unterschiede in der Testkonzeption mathematischer Kompetenz in PISA und den Bildungsstandards in Deutschland ermöglicht eine weitergehende Interpretation der Befunde aus PISA 2012, die auch für die Verwendung der Ergebnisse im Kontext der deutschen Bildungspolitik relevant ist. Beispielsweise wird deutlich, dass die signifikante Verbesserung der durchschnittlichen mathematischen Kompetenz der Schülerinnen und Schüler in Deutschland seit PISA 2003 vermutlich insbesondere auf Schülerinnen und Schüler an nicht gymnasialen Schulformen zurückgeht. Dieser Aspekt wird in Beitrag 3 kritisch diskutiert, da die Abbildung von Trends in verschiedenen Schularten durch die erhebliche Veränderung der Zusammensetzung der Schülerinnen und Schüler in den Schularten seit PISA 2000 kaum möglich und auch bei den in ihrer Grundkonzeption weitgehend unveränderten Gymnasien (Cortina et al., 2003; Neugebauer, 2010) aufgrund der gestiegenen Bildungsbeteiligung mit großer Vorsicht zu interpretieren ist. Es wird argumentiert, dass zur über die fünf bisher abgeschlossenen PISA-Erhebungsrunden festgestellten Verbesserung der Schülerleistungen insbesondere in den nicht gymnasialen Schulformen die 2003 und 2004 eingeführten Bildungsstandards in Mathematik für den Hauptschul- und den Mittleren Schulabschluss beigetragen haben könnten. Durch den klaren Bezug der PISA-Ergebnisse zum ebenfalls 2012 durchgeführten IQB-Ländervergleich zur Überprüfung der Bildungsstandards, der nach Abschluss des ersten PISA-Zyklus (2000-2006) aus der PISA-Erweiterung (PISA-E) zum Vergleich der Bundesländer hervorging, wird herausgearbeitet, wie sich beide Schulleistungsstudien auf der Sekundarstufe I ergänzen und nicht – wie oftmals irrtümlich angenommen – austauschbar sind.

In **Beitrag 4** (Hohn et al., 2013) wird mit der Lesekompetenz eine weitere in PISA erfasste Domäne mit ihrem Testkonzept und aktuellen Ergebnissen aufgegriffen. In diesem Beitrag war zu untersuchen, wie sich die Lesekompetenz seit PISA 2000 bei den Schülerinnen und Schülern in Deutschland entwickelt hat und inwieweit den zahlreichen in den ersten PISA-Ergebnissen gefundenen Disparitäten seither nachgegangen worden ist. Zu dem auf die Ergebnisse aus PISA 2000 folgenden PISA-Schock (Roeder, 2003) hatten neben dem unterdurchschnittlichen Niveau der Lesekompetenz deutscher Schülerinnen und Schüler auch enorme Disparitäten innerhalb Deutschlands beigetragen (Baumert et al., 2001; Baumert et al., 2002). Regionale Unterschiede, Differenzen in den Leseleistungen zwischen Mädchen und Jungen sowie zwischen Jugendlichen mit einem Zuwanderungshintergrund und solchen ohne gaben Anlass zur Sorge und wurden unter anderem in den sieben zentralen Handlungsfeldern der KMK (KMK, 2002) auf die Agenda der Bildungspolitik gesetzt. Gezielte Förderung der Lesekompetenz, insbesondere für Jugendliche mit einem Zuwanderungshintergrund, ist ein Kernthema dieser Handlungsfelder. Entsprechend wurde die 2012 erstmals signifikant über dem OECD-Durchschnitt liegende Lesekompetenz der Schülerinnen und Schüler in Deutschland in Beitrag 4 in den Kontext verschiedener bildungspolitischer Anstrengungen zur Förderung der Lesekompetenz eingeordnet, aber auch kritisch diskutiert. Wie auch in Beitrag 3 wurde die Abbildung von Trends in verschiedenen Schularten seit PISA 2000 sehr vorsichtig interpretiert. Gleichwohl ist ein bemerkenswerter Befund des vierten Beitrags, dass die seit 2000 von 28 auf 36 Prozent gestiegene Bildungsbeteiligung an Gymnasien in

Deutschland zu keinem Niveauverlust bei der Leseleistung an dieser Schulform geführt hat. Im Durchschnitt lesen die Gymnasiastinnen und Gymnasiasten genauso gut wie im Jahr 2000 (582 Punkte, weit über dem deutschen Gesamtdurchschnitt; vgl. Artelt et al., 2001). Die Analysen zeigten weiterhin, dass 2012 deutlich mehr Schülerinnen und Schüler entsprechend gut lesen. Ähnlich wie bei der mathematischen Kompetenz haben sich die Leistungsunterschiede zwischen Mädchen und Jungen über die Zeit nicht verringert. Zwar ist bei beiden Geschlechtern ein kontinuierlicher Aufwärtstrend erkennbar, aber die Leistungen der beiden Gruppen nähern sich einander nicht an.

Während die Beiträge 3 und 4 Kompetenz auf der Basis von Tests analysiert haben, legt **Beitrag 5** (Sälzer et al., 2013) einen Schwerpunkt auf die schulischen Rahmenbedingungen für die Entwicklung der in PISA gemessenen Kompetenzen. Die theoretische Einbettung der Thematik basiert auf der Unterscheidung der Schule als Institution, etwa mit Blick auf rechtliche und administrative Rahmenbedingungen, von der Einzelschule als Ort des Lernens und Lebensraum mit bestimmten Merkmalen (van Ackeren et al., 2015). Schule wird durch diese Unterscheidung neben ihrer Qualifikationsfunktion als Vorbereitung für die Anforderungen in Beruf, Privatleben und Gesellschaft auch als Sozialisationsinstanz aufgefasst (Blömeke & Herzig, 2009; Fend, 1981; Fend, 2006). Anders als die Beiträge 3 und 4, die sich der Beschreibung von Schülerkompetenzen und damit dem Monitoring widmen, stellt Beitrag 5 die Frage, welche Merkmale des deutschen Schulsystems und deutscher Schulen im Vergleich zu anderen Bildungssystemen auffällig sind (Benchmarking). Im Unterschied zur Auswertungsstrategie der OECD ist nicht die statistische Erklärung von Leistungsvarianz durch Schulmerkmale zentral (OECD, 2013b), sondern eine theoretische Interpretation der Bedeutung von Indikatoren auf der Systemebene, die einen Bezug sowohl zu strukturellen Bedingungen als auch zu Einzelschulen in Deutschland haben und insbesondere vor dem Hintergrund einiger Besonderheiten des deutschen Schulsystems eingeordnet werden. Der in Beitrag 5 mit der Aufteilung der Varianz der untersuchten Variablen (hier: mathematische Kompetenz) in Komponenten innerhalb von Schulen und zwischen den Schulen im ersten Schritt berichtete zentrale Kennwert, der in Large-Scale Assessments zu den grundlegenden Indikatoren gehört (vgl. u. a. Fay, 1989; Martin & Mullis, 2012; Rust & Rao, 1996), wird bezüglich der Grenzen seiner Aussagekraft kritisch reflektiert. Diese Aufschlüsselung der Varianz kann Anhaltspunkte über Qualitätsunterschiede zwischen Schulen geben, aber auch zu unterschiedlichen Zusammensetzungen und Rekrutierungsstrategien von Schulen. Allerdings reichen die durch PISA gewonnenen Daten nicht aus, um spezifisch Schulentwicklungsprozesse auf der Ebene der Einzelschule zu begründen. Die Daten können jedoch beispielsweise durch längsschnittliche Erweiterungen entsprechend angereichert werden (vgl. etwa Bischof et al., 2013). Im zweiten Schritt stellt Beitrag 5 Ergebnisse aus Analysen zu den beiden genannten Ebenen der schulischen Rahmenbedingungen für die Entwicklung von Kompetenzen dar, indem zunächst auf institutionelle und organisatorische Rahmenbedingungen und im Anschluss auf Merkmale von Einzelschulen eingegangen wird. Der Beitrag greift ausgewählte Indikatoren über mehrere PISA-Erhebungsrunden auf, deren Bedeutung ohne internationale Vergleichsstudien wie PISA womöglich nach wie vor unklar wäre. Dazu

gehören etwa das Alter der Schülerinnen und Schüler bei der Differenzierung in verschiedene Schulformen der Sekundarstufe, der Anteil der Klassenwiederholungen bis zum Alter von 15 Jahren oder die Anzahl der verfügbaren Schulformen auf der Sekundarstufe (vgl. u. a. Baumert et al., 2001; Fuchs, 2003). Insgesamt zeigt der Beitrag, dass die Schulen in Deutschland im internationalen Vergleich wenige Problemlagen (z. B. Beeinträchtigung des Lernens zur schwänzende oder zu spät kommende Schüler) aufweisen. In Bezug auf Maßnahmen zur Qualitätssicherung an Schulen wird herausgearbeitet, dass Schulen in Deutschland einen Teil der Instrumente der Qualitätssicherung und –entwicklung nutzen, die auch in anderen Bildungssystemen üblich sind. Die Begleitung von Experten für Schulentwicklung beispielsweise wird in Deutschland im Vergleich zu anderen OECD-Staaten nur selten in Anspruch genommen. Der Text hebt aber auch hervor, dass aufgrund des Designs von PISA weder das mittlere Leistungsniveau der Jugendlichen eines Staates noch der nationale Leistungstrend durch die Strukturmerkmale erklärt werden kann.

**Vertiefung des Kenntnisstandes im Bereich internationales Bildungsmonitoring.** Large-Scale Assessments verbinden im Forschungsfeld der empirischen Bildungsforschung Fragen nach strukturellen Merkmalen des Bildungswesens wie der Qualität und Quantität von Ressourcen mit Fragen zu Ergebnissen, Veränderungen und Folgen von Bildungsprozessen (vgl. etwa Straus, 2014; Stufflebeam, 1983). Eine der zentralen Funktionen von regelmäßig durchgeführten Large-Scale Assessments ist die Beobachtung von Entwicklungen über die Zeit, die als Grundlage sowohl für die Identifikation von Schwachstellen dienen kann als auch für das Monitoring von Veränderungen. In diesem Kontext nimmt **Beitrag 6** eine systematisch reflektierende Perspektive ein und verfolgt die Fragestellung, welche Faktoren zu der im internationalen Vergleich einzigartig positiven Entwicklung Deutschlands in Bezug auf die durchschnittlichen Schülerkompetenzen in PISA zwischen 2000 und 2012 beigetragen haben könnten. In einem ersten Schritt werden einige zentrale Ergebnisse aus PISA 2000 in Erinnerung gerufen und ausgewählten Befunden aus PISA 2012 gegenübergestellt. Auf diese Weise wird deutlich, wie unterschiedlich die Gesamtergebnisse aus PISA 2000 und 2012 sind und weshalb das überdurchschnittliche mittlere Kompetenzniveau in PISA 2012 in Deutschland als besonders erfreulich und keinesfalls selbstverständlich gelten kann. Im nächsten Schritt werden drei Aspekte eines bildungspolitisch und wissenschaftlich gesteuerten Veränderungsprozesses hervorgehoben: Eine gründliche Diagnose der Ausgangslage, ein intensiver Diskurs zwischen Politik, Öffentlichkeit und Wissenschaft sowie einige Beispiele für Maßnahmen, die nach PISA 2000 bundesweit und damit länderübergreifend ergriffen wurden, um die Voraussetzungen für den Kompetenzerwerb an deutschen Schulen nachhaltig zu verbessern. Damit geht der Beitrag auf eine aktuelle Debatte in der Bildungsforschung ein, in welcher diskutiert wird, ob Large-Scale Assessments eine Vereinfachung und Reduktion des Bildungsbegriffs auf messbare, anschlussfähige Kompetenzen befördern (Berliner, 2011; Carnoy, 2014). In Beitrag 6 wird argumentiert, dass PISA als Large-Scale Assessment keineswegs einen Bildungsbegriff reduziert und durchaus wesentliche Merkmale von Schülerinnen und Schülern und ihrer Lernumgebung neben den durchschnittlichen Kompetenzen

erfasst. Dabei ist der Kern des Nutzens von Daten aus PISA jedoch der, diese innerhalb ihrer Möglichkeiten und vor allem auch Grenzen zu interpretieren (vgl. auch Baumert und Köller, 1998). PISA liefert als Bezugspunkt für eine Diagnose des jeweiligen Ist-Zustandes Indikatoren für einen empirisch fundierten Diskurs zwischen verschiedenen Akteuren bildungspolitischer Steuerung. Diese Indikatoren sind von hoher Qualität und können wichtige Hinweise darauf liefern, wo Strukturen, Prozesse oder Ergebnisse im deutschen Bildungswesen verbesserungswürdig sind. Die PISA-Daten reichen jedoch nicht darüber hinaus. In diesem Sinne werden abschließend die übergreifenden, bundesweiten Maßnahmen zur Verbesserung der Rahmenbedingungen für die Kompetenzentwicklung diskutiert, durch deren Evaluation durchaus mehr und andere empirische Evidenz gewonnen werden konnte als aus PISA allein (vgl. etwa Ostermeier et al., 2010; Prenzel et al., 2009).

Deutschland beteiligt sich mit PISA, TIMSS und IGLU/PIRLS aktuell an drei regelmäßig stattfindenden international vergleichenden Schulleistungsstudien. Während in PISA mit der Population fünfzehnjähriger Schülerinnen und Schüler die Sekundarstufe I im Fokus steht, handelt es sich bei TIMSS und IGLU um Grundschuluntersuchungen. Grundsätzlich liefern Befunde aus Large-Scale Assessments Anhaltspunkte dafür, wo ein Bildungssystem im internationalen Vergleich steht und wo sich möglicherweise Hinweise auf Herausforderungen und Potenzial für Verbesserungen finden lassen (Bromme et al., 2014; Drechsel & Prenzel, 2008; Drechsel et al., 2009). Dabei stellen die Ergebnisse einer einzelnen Erhebungsrunde solcher Untersuchungen, etwa in PISA, eine Momentaufnahme dar. Ein deutlicher Mehrwert von Befunden aus Studien wie PISA liegt darin, die Resultate mehrerer Erhebungsrunden gemeinsam zu betrachten und darin erkennbare Entwicklungen in Bezug auf den Bildungs- und Systemkontext zu interpretieren; dies insbesondere dann, wenn mehrere Studien gemeinsam betrachtet werden. PISA hat seit der ersten Erhebungsrunde im Jahr 2000 ebenso wie TIMSS drei Jahre zuvor und IGLU seit 2001 enorm dazu beigetragen, dass Bildung in Deutschland einerseits wieder ein wichtiges Thema öffentlicher Diskussion geworden ist (Artelt et al., 2007; Stanat & Pant, 2014) und andererseits einen Wechsel von der Input- hin zur Outputsteuerung angeregt (Klieme & Vieluf, 2013). In diesem Sinne werden Bildungsprozesse von ihren Ergebnissen her gedacht und reflektiert, so dass beispielsweise die Beherrschung von Mindestkompetenzen als Ziel von Unterricht definiert werden kann (vgl. etwa Blum et al., 2010; Tenorth, 2004). Als international vergleichende Studie zum Bildungsmonitoring und –benchmarking hat PISA darüber hinaus eine Perspektive auf Bildung eingenommen, die zuvor über lange Zeit im Kontext erziehungswissenschaftlicher Forschung in Deutschland nicht vorhanden gewesen war: Bildungsergebnisse und die die Untersuchung des Zusammenhangs mit ihren Rahmenbedingungen (vgl. etwa Gogolin, 2010; Tillmann et al., 2008). **Beitrag 7** (Sälzer et al., in Druck) bezieht sich auf diesen fachwissenschaftlichen Diskurs, indem anhand der drei aktuellen international vergleichenden Schulleistungsstudien mit deutscher Beteiligung eine bilanzierende Betrachtung der jeweiligen Ergebnisse aus mehreren Erhebungsrunden vorgenommen wird. Theoretisch baut der Beitrag auf der von Rolff (2000) vorgeschlagenen Trias der Schulentwicklung in Form von Unterrichts-, Personal- sowie Organisationsentwicklung auf. Die Frage,

inwieweit internationale Vergleichsstudien einen Beitrag zur Schulentwicklung leisten können, steht im Fokus des Beitrags. Mit dem im ersten Schritt beschriebenen Aufbau und Ablauf der Studien sowie einer Beschreibung der darin gemessenen Kompetenzen wird zunächst die Rückmeldefunktion (Kuper, 2006) des internationalen Vergleichs von Schülerleistungen herausgearbeitet. Im nächsten Schritt folgt eine verdichtete Darstellung der zentralen Ergebnisse aus TIMSS, IGLU und PISA über die bisher abgeschlossenen Erhebungsrounds hinweg. Alle drei Studien beschränken sich nicht auf die Erhebung kognitiver Indikatoren, sondern erfassen darüber hinaus auch Einstellungen, Merkmale und Kontextfaktoren der Schülerinnen und Schüler und ihres schulischen und häuslichen Umfelds (vgl. Mullis et al., 2012; Mullis et al., 2009; OECD, 2013a). Damit sind sie als Instrument für die Untersuchung des Zusammenhangs zwischen Rahmenbedingungen für Bildung und Kompetenzerwerb und entsprechende Ergebnisse adäquat und können zugleich ein Diagnoseinstrument für das Erreichen mehrdimensionaler Bildungsziele in Schulen sein. Für alle drei von Rolff (2000) genannten Dimensionen der Schulentwicklung lassen sich im abschließenden Schritt des Beitrags Ansatzpunkte in den Ergebnissen der beschriebenen Vergleichsstudien finden. Als Beispiel für Unterrichts- und Personalentwicklung können unter anderem Befunde zum Bedarf an verstärkter Talentförderung dienen; das Thema Organisationentwicklung wird etwa in der Analyse von Schulprogrammen und deren Wirkung auf Schülerkompetenzen bearbeitet.

### ***Forschungsbereich 2:***

#### ***Differenzierende Befunde aus Large-Scale Assessments für Teilpopulationen von Schülerinnen und Schülern***

Eine wichtige Funktion von Large-Scale Assessments ist die Gewinnung von beschreibendem Wissen über Kompetenzen von Schülerinnen und Schülern. International vergleichende Large-Scale Assessments stellen dazu Indikatoren auf repräsentativer Basis bereit, die Erkenntnisse auf verschiedenen Ebenen des Bildungssystems ermöglichen (Klieme, 2013; Prenzel, 2005). Der Anspruch der Repräsentativität vergleichender Schulleistungsstudien bringt mit sich, dass einzelne Teilpopulationen von Schülerinnen und Schülern auf dieser Basis nicht vertieft untersucht werden können, da ihr Anteil an der untersuchten Population meist zu klein für differenzierte Datenauswertungen ist (Blanck et al., 2013). Dafür sind üblicherweise Ergänzungsstichproben notwendig. Abhängig vom Design zur Stichprobenziehung und entsprechender Stratifizierung ist es jedoch beispielsweise möglich, Unterschiede zwischen Mädchen und Jungen oder zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund repräsentativ zu untersuchen, so dass auf die jeweilige Zielpopulation der Studie geschlossen werden kann. Allerdings sind unter diesen Teilpopulationen meist weitere Subgruppen zu finden, für die anhand der gezogenen Stichprobe keine gültigen Aussagen getroffen werden können (z. B. Ethnien innerhalb der Gruppe der Schülerinnen und Schüler mit Zuwanderungshintergrund). Dementsprechend liegen derzeit auch nur wenige Arbeiten mit repräsentativen Stichproben zu Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in

Deutschland vor (vgl. etwa Kocaj et al., 2014), da ihr Anteil an der Zielpopulation gering ist und dementsprechend auch die Stichprobe dieser Schülerinnen und Schüler keine verallgemeinerbaren Aussagen zulässt. Auch die theoretische und methodische Basis innerhalb der empirischen Bildungsforschung in Bezug auf die Berücksichtigung von Schülerinnen und Schülern mit besonderem Förderbedarf in vergleichenden Schulleistungsstudien befindet sich derzeit noch in einem relativ wenig entwickelten Stadium (Pohl & Carstensen, 2013). Gleichwohl sind Erkenntnisse über durchschnittliche Kompetenzen von Schülerinnen und Schülern mit besonderem Förderbedarf sowie deren Lern- und Entwicklungsbedingungen in verschiedenen schulischen Lernsettings von zentraler Bedeutung für das Bildungsmonitoring (Heydrich et al., 2013; Kuhl et al., 2015). Die Beiträge, die für den zweiten Forschungsbereich in die Habilitation eingehen sollen, greifen diesen Bedarf auf und erweitern die nationale Berichterstattung zu PISA 2012 um Analysen zu einer Teilpopulation von Schülerinnen und Schülern, die anhand einer Ergänzungsstichprobe von 49 Sonder- und Förderschulen zu PISA 2012 deutschlandweit untersucht werden konnte: Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf.

**Untersuchung von Teilpopulationen in PISA.** PISA ist als international angelegte Bildungsstudie darauf ausgerichtet, repräsentative Stichproben von Schülerinnen und Schülern *verschiedener Staaten* auf nationaler Ebene miteinander zu vergleichen. Diese Vergleichbarkeit muss durch zahlreiche Konsentscheidungen, etwa über die Definition von Kompetenzen oder Kategorien von Schülermerkmalen, abgesichert werden. In Bezug auf die Erfassung, Untersuchung und Beschreibung spezieller Teilpopulationen hat dies verschiedene Konsequenzen. Der zweite Forschungsbereich der vorliegenden Habilitationsschrift greift den Anspruch an Schulleistungsvergleichsstudien zur Bereitstellung möglichst repräsentativer Indikatoren auf und nimmt eine Teilgruppe der PISA-Zielpopulation fünfzehnjähriger Schülerinnen und Schüler in Deutschland in den Blick. Im Forschungsfeld der empirischen Bildungsforschung werden aktuell mehrere Herausforderungen diskutiert, die mit dem Einbezug von Schülerinnen und Schülern mit besonderem Förderbedarf in Large-Scale Assessments verbunden sind (vgl. etwa Kuhl et al., 2015; Pohl & Carstensen, 2013). Diese Herausforderungen sind einerseits methodischer Art in Bezug auf die Validität und Reliabilität der Testinstrumente (Heydrich et al., 2013) oder deren mögliche Anpassung an die Bedürfnisse der getesteten Schülerinnen und Schüler (Nusser et al., 2015) und andererseits konzeptueller Art hinsichtlich der Definition sonderpädagogischen Förderbedarfs sowie entsprechender Zuweisung von Kategorien des Förderschwerpunktes (Gresch & Piezunka, 2015). Auf internationaler Ebene ist mit der Verabschiedung und weitgehenden Ratifizierung der UN-Behindertenrechtskonvention im Jahr 2006 bzw. 2009 zwar eine gemeinsame Grundlage für die Zuweisung eines Förderbedarfs sowie den Einbezug von Schülerinnen und Schülern mit besonderem Förderbedarf in vielen Schulsystemen gelegt worden, diese beeinflusst jedoch nicht die nationalen und teils föderalen Kriterien und Definitionen für die Zuschreibung eines sonderpädagogischen Förderbedarfs. Aktuell fehlt eine systematische Aufbereitung

der Kriterien, Definitionen und Zuweisungsprozesse eines Förderbedarfs für Deutschland, was wiederum die Inklusion dieser Teilpopulation in Large-Scale Assessments erschwert.

Dieser konzeptuellen Problemstellung geht **Beitrag 8** nach. Den Ausgangspunkt des Beitrags bilden in erster Linie rechtliche Grundlagen wie die allgemeine Schulpflicht für Menschen mit Behinderung und deren Regelung in den Landesverfassungen (Pestalozza, 2014). Die allgemeine Schulpflicht ist in Deutschland aufgrund der Kulturhoheit der Bundesländer in deren Landesverfassungen festgelegt. Sie gilt zunächst grundsätzlich auch für Kinder mit einer Behinderung bzw. mit sonderpädagogischem Förderbedarf. Zwar ist die institutionelle Organisation der sonderpädagogischen Förderung bundeseinheitlich geregelt (KMK, 1994; KMK, 2011), die Ausgestaltung obliegt jedoch den einzelnen Bundesländern (Bundschuh & Baier, 2009). Beitrag 8 verfolgt die Fragestellung, wie der Prozess der Feststellung eines sonderpädagogischen Förderbedarfs in den deutschen Bundesländern jeweils verläuft und welche Akteure daran beteiligt sind. Unterschieden werden dabei der Anstoß des Diagnoseprozesses, die Prozessverantwortung, die Schulwahl sowie die letztliche Entscheidung über den Beschulungsort in Konfliktfällen. Auf diese Weise werden Schnittstellen im Prozess identifiziert und systematisch dargestellt. Der Beitrag greift damit in den aktuellen Diskurs in der empirischen Bildungsforschung ein und liefert eine konzeptuelle Grundlage für die Definition eines sonderpädagogischen Förderbedarfs in bundesweiten Schulleistungsstudien, indem die Prozesse und Etikettierungen in allen 16 Bundesländern dargestellt und beschrieben werden. Dies trägt dem Bedarf Rechnung, die Beteiligung von Schülerinnen und Schülern mit Förderbedarf an Schulleistungsvergleichsstudien zu ermöglichen und ihre Berücksichtigung theoretisch abzusichern. In dem Beitrag wird zunächst differenziert, wann der Prozess der Feststellung eines sonderpädagogischen Förderbedarfs beginnt; dies kann sowohl vor als auch nach der Einschulung geschehen. In jedem Fall ist die professionelle Diagnose eines sonderpädagogischen Förderbedarfs Voraussetzung für dessen offizielle Zuweisung (Bundschuh, 2006). Trotz einer gemeinsamen Grundlage durch die KMK zeigt Beitrag 8, dass wesentliche Schnittstellen im Prozess der Feststellung eines sonderpädagogischen Förderbedarfs in Deutschland zwischen den Bundesländern variieren. So kann der Anstoß des Diagnoseprozesses durch Eltern, Schule, Schularzt oder eine vorschulische Einrichtung erfolgen. Die Prozessverantwortung liegt je nach Bundesland beim staatlichen Schulamt, einer Behörde der Schulaufsicht, der Schulleitung, bei einem Förderzentrum oder einem mobilen sonderpädagogischen Dienst. Auch die Schulwahl liegt in unterschiedlichen Bundesländern in der Zuständigkeit verschiedener Personen. Dies können die Eltern bzw. Erziehungsberechtigten sein, aber auch ein Förderausschuss, eine Bildungswegekonferenz, die Schulleitung oder eine Behörde der Schulaufsicht. In allen Bundesländern ist darüber hinaus geregelt, wer die letztgültige Entscheidung über den Beschulungsort trifft. Dies sind in keinem Fall die Eltern, sondern stets Behörden. Wichtigstes Fazit des Beitrages ist, dass die aktuelle Situation mit 16 unterschiedlichen Wegen zur Zuweisung eines sonderpädagogischen Förderbedarfs und damit verbundener Schulwahl bedeutet, dass es letzten Endes darauf ankommt, in welchem Bundesland ein Kind aufwächst. Eine Einigung auf ein gemeinsames,

bundesweites Vorgehen auf institutioneller Ebene käme auch den nicht-institutionellen Akteuren, nämlich den Eltern und ihrem Kind, zu Gute. Einerseits folgte die Förderdiagnostik einem gemeinsamen Modell und die Entscheidung über die geeignetste Art der Beschulung würde erleichtert.

Die aktuellen Herausforderungen in der empirischen Bildungsforschung hinsichtlich der Beschreibung der Kompetenzen von Schülerinnen und Schülern mit besonderem Förderbedarf sind neben den genannten konzeptuellen bzw. theoretischen Problemen auch methodischer Art. So sind etwa schulartspezifische Aussagen für Schülerinnen und Schüler an Förderschulen anhand der international üblicherweise erhobenen PISA-Daten aufgrund des geringen Anteils an dieser Schülerinnen und Schüler an der Gesamtpopulation nicht zulässig (Sälzer & Prenzel, 2013). Im Zuge der UN-Behindertenrechtskonvention und der damit verbundenen Forderung nach inklusiver Beschulung (United Nations Office of Legal Affairs, 2006) sowie innerhalb des Fachdiskurses in der empirischen Bildungsforschung wurde in den letzten Jahren der Bedarf an belastbaren Vergleichsdaten zu den Kompetenzen und schulischen Rahmenbedingungen von Schülerinnen und Schülern mit Förderbedarf deutlich artikuliert (Blanck et al., 2013). Aus diesem Grund wurde in Begleitung zu PISA 2012 eine zusätzliche Stichprobe von 49 Förderschulen gezogen, die gemäß den PISA-Prozeduren getestet wurde. Die **Beiträge 9** und **10** berichten empirische Ergebnisse aus diesem Förderschul-*Oversample*. Beide Aufsätze nutzen sowohl die Schülerinnen und Schüler aus der Erweiterungsstichprobe als auch die Schülerinnen und Schüler mit Förderbedarf, die Teil der regulären PISA-Stichprobe Deutschlands bzw. der Neuntklässlerstichprobe in PISA 2012 sind. Die Neuntklässlerstichprobe ist ebenfalls eine Erweiterung des internationalen Pflichtprogramms und dient dazu, geteilte Wahrnehmungen und Unterrichtserfahrungen von Schülerinnen und Schülern zu erfassen.

**Beitrag 9** (Gebhardt et al., 2015) legt besonderes Augenmerk auf einen Vergleich der mathematischen Kompetenz zwischen Schülerinnen und Schülern an Förderschulen und solchen in inklusiven Unterrichtsettings. Die Ausgangsfrage lautet, ob sich die durchschnittlichen Kompetenzen von Schülerinnen und Schülern mit Förderbedarf an Förderschulen systematisch von solchen unterscheiden, die in inklusiven Settings unterrichtet werden. Erste empirische Befunde deuten darauf hin, dass sich beide Gruppen systematisch voneinander unterscheiden, beziehen sich allerdings bisher ausschließlich auf die Grundschulstufe (Kocaj et al., 2014; Kocaj et al., 2015). Die theoretische Basis dieses Beitrags bilden neben der Definition der untersuchten Gruppe von Schülern mit Förderschwerpunkt Lernen mehrere internationale Befunde, die sowohl der Abgrenzung dieses Förderschwerpunktes gegenüber weiteren Ursachen für stark unterdurchschnittliche Schulleistungen als auch der Klärung des Begriffes schulischer Inklusion dienen (Baker et al., 1995; Myklebust, 2002). Darauf aufbauend werden Schülerinnen und Schüler mit dem Förderschwerpunkt Lernen in zwei Gruppen miteinander verglichen: solche, die in Förderschulen unterrichtet werden und solche, die an Regelschulen und damit in inklusiven Settings beschult werden. Beide Gruppen finden sehr unterschiedliche Rahmenbedingungen für ihren Kompetenzerwerb vor, die sich beispielsweise in der Klassengröße, der Ausstattung der Schule

mit Fachpersonal für sonderpädagogische Förderung oder der Art des Curriculums ausdrücken (vgl. etwa Blanck, 2014; Lüke und Ritterfeld, 2012; Schuck, 2014). Das zentrale Ergebnis der Studie ist, dass die Jugendlichen in inklusiven Lernumgebungen im PISA-Test im Durchschnitt eine Kompetenzstufe (II) besser abschneiden als Schülerinnen und Schüler an Förderschulen (Stufe I). Obwohl beide Gruppen am unteren Ende des Kompetenzspektrums sind – oder gerade deswegen, da eine geringe Varianz vorliegt – ist die Deutlichkeit dieses Befunds bemerkenswert. Darüber hinaus zeigt sich, dass Schülerinnen und Schüler mit Förderbedarf im Mittel einen geringeren sozioökonomischen Status aufweisen als ihre Klassenkameraden ohne Förderbedarf. Dies gilt sowohl für die Jugendlichen an Förderschulen als auch für jene in inklusiven Klassen. Die Befunde des Beitrags wurden auch in anderen Studien, unter anderem mit Daten des IQB-Ländervergleichs 2011 für die Primarstufe, bestätigt (Kocaj et al., 2014; Kocaj et al., 2015). Mögliche Erklärungen für das bessere Abschneiden der Schüler in inklusiven Klassen sind, dass diese unter Umständen von der Interaktion mit leistungsstärkeren Klassenkameraden profitieren oder dass tendenziell diejenigen Schülerinnen und Schüler, bei denen eine vergleichsweise leichte Form der Behinderung vorliegt, eher in inklusiven Settings beschult werden als Schülerinnen und Schüler mit einer ausgeprägteren Behinderung (Myklebust, 2002).

**Beitrag 10** knüpft an die Ergebnisse aus Beitrag 9 an und beschreibt die durchschnittlichen Kompetenzen in Mathematik, Lesen und Naturwissenschaften für die Schülerinnen und Schüler mit besonderem Förderbedarf aller untersuchten Förderschwerpunkte an den 49 Förderschulen der erweiterten Stichprobe sowie in der regulären PISA-Stichprobe 2012. Der Beitrag lässt sich daher in eine bis dahin bestehende Forschungslücke bezüglich der Beschreibung der bundesweiten durchschnittlichen Leistungen in PISA und damit auf der Sekundarstufe I einordnen, die zuvor noch nicht existierte. Die Studie geht damit über bislang vorliegende, aber regional begrenzte Befunde hinaus (z. B. Bos et al., 2009; Lehmann & Hoffmann, 2009). Im Zentrum des Beitrags steht ein Vergleich der mittleren Kompetenzen von Fünfzehnjährigen mit besonderem Förderbedarf an Förderschulen und Fünfzehnjährigen ohne Förderbedarf an den übrigen allgemeinbildenden Schularten der Sekundarstufe. Erwartungsgemäß ist das Kompetenzniveau der untersuchten Schülerinnen und Schüler an Förderschulen niedriger als das der Fünfzehnjährigen an anderen Schularten, allerdings treten einige förderschwerpunkt- sowie geschlechtsspezifische Unterschiede hervor. Beispielsweise schneiden Schülerinnen und Schüler, die dem Förderschwerpunkt emotionale und soziale Entwicklung zugeordnet sind, signifikant besser ab als Jugendliche mit dem Förderschwerpunkt Lernen. Während in den Domänen Mathematik und Lesen wie auch in der regulären PISA-Stichprobe Kompetenzunterschiede zu Gunsten der Jungen (Mathematik) bzw. der Mädchen (Lesen) gefunden werden, so erzielen die Jungen unter den Förderschülern auch in den Naturwissenschaften durchschnittlich bessere Leistungen als die Mädchen. In der regulären PISA-Stichprobe liegt keine solche Disparität vor (Schiepe-Tiska et al., 2013). Der Beitrag befindet sich aktuell in der finalen Überarbeitungsphase für die Zeitschrift Unterrichtswissenschaft und wird voraussichtlich 2016 erscheinen.

Die zehn Beiträge, die in die vorgelegte Habilitationsschrift eingehen, haben als gemeinsamen Bezugspunkt die Befunde aus verschiedenen PISA-Erhebungsrunden, insbesondere PISA 2012. PISA steht dabei als Large-Scale Assessment in seiner Funktion für ein Bildungsmonitoring im Mittelpunkt. Während sich der erste Forschungsbereich der Habilitation der Rolle von Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen widmet, knüpft der zweite Forschungsbereich daran an und fügt Befunde aus Large-Scale Assessments für bestimmte Teilpopulationen von Schülerinnen und Schülern hinzu. Die Berücksichtigung beispielsweise von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf schränkt möglicherweise die internationale Vergleichbarkeit von Large-Scale Assessments ein und muss aus diesem Grund besonders gut konzeptuell und methodisch abgesichert werden. Einschränkungen können etwa aufgrund der in verschiedenen Bildungssystemen unterschiedlichen Kriterien für die Zuweisung eines sonderpädagogischen Förderbedarfs entstehen oder in Bezug auf die Definition eines Förderbedarfs mit oder ohne formale Diagnostik. Zwar sind genaue Samplingvorschriften ohnehin unerlässlich für einen belastbaren Vergleich, aber für Aussagen über Teilpopulationen werden Ergänzungsstichproben (*Oversamples*) benötigt.

Im Rahmen der Beiträge geklärt werden konnten insbesondere Fragen zum Kompetenzniveau der Jugendlichen in Deutschland, Entwicklungen der Schülerkompetenzen über mehrere Erhebungsrunden sowie verbleibende Herausforderungen an das deutsche Bildungssystem und die Bildungspraxis, um positive Veränderungen stabilisieren und weiterführen zu können. Darüber hinaus konnten Aspekte der Nutzung von Ergebnissen aus Large-Scale Assessments ebenso wie theoretische Befunde zur Zuweisung von sonderpädagogischem Förderbedarf und einige empirische Ergebnisse für Schülerinnen und Schüler an Förderschulen herausgearbeitet werden.

Large-Scale Assessments werden in den vorliegenden Beiträgen in das Forschungsfeld der empirischen Bildungsforschung eingeordnet, indem zumindest Voraussetzungen und Ergebnisse von Bildung analysiert, beschrieben und interpretiert werden. Der Aspekt der Bildungsprozesse, der ebenfalls zu den Hauptanliegen empirischer Bildungsforschung gehört (Gräsel, 2011; Prenzel, 2005), ist lediglich in Form der Trendbetrachtungen von PISA-Befunden über mehrere Erhebungsrunden hinweg betrachtet worden und nicht im Sinne einer Verlaufsanalyse. Dies ist weitestgehend dem Design der PISA-Studie geschuldet, die in Form der Daten aus PISA 2012 sowie der Stichprobenerweiterung der Förderschulen lediglich einen Messzeitpunkt erfasst und demnach nicht geeignet ist für die Analyse von Prozessen. Vielmehr fokussieren die gewählten Fragestellungen und Auswertungen in den Beiträgen zu beiden Forschungsbereichen eine Beschreibung von Kompetenzen, deren Entwicklung über die Zeit und deren Entstehungsbedingungen. Zum Verhältnis von Bildungsforschung zu Bildungspolitik liegen allerdings bereits zahlreiche einschlägige Arbeiten vor, die im Fachdiskurs der empirischen Bildungsforschung rezipiert worden sind (z. B. Bellmann, 2006; Bromme et al., 2014; Terhart, 2001; Tillmann & Vollstädt, 2001). Insofern schließt die vorliegende Habilitation einige Forschungslücken im Feld der empirischen

Bildungsforschung, die inhaltlich Gegenstand bildungspolitischer Steuerung sind, ohne deren Rezeption und Verarbeitung durch bildungspolitische Akteure zu thematisieren.

Zusammengefasst werden vergleichende Schulleistungsstudien in den vorliegenden Arbeiten in ihrer Funktion als Bildungsmonitoring und Benchmarking genutzt, um Wissen über den Bildungsstand Jugendlicher in Deutschland und damit verbundene Rahmenbedingungen zu generieren. Auf diese Weise konnte der Kenntnisstand im Bereich internationales Bildungsmonitoring vertieft werden. Unter den vorgelegten Aufsätzen ist auch ein Beitrag, der erörtert, inwieweit ein Bezug zwischen den Erkenntnissen aus international vergleichenden Schulleistungsstudien und Schulentwicklung in Deutschland hergestellt werden kann. Dieser Beitrag bestätigt, dass die aus Large-Scale Assessments gewonnene Datengrundlage nicht für die unmittelbare Ableitung von Schulentwicklungsprozessen geeignet ist. Um Schulentwicklung anzustoßen, zu begleiten und zu evaluieren, sind andere Ansätze des Erkenntnisgewinns notwendig (Fickermann & Maritzen, 2014).

### **3. Eigenleistungen im Kontext der Fachpublikationen**

Die Fachpublikationen, die in die Habilitation eingehen, lassen sich im Kontext der PISA-Studie verorten. Dennoch hat jeder Beitrag einen eigenen Entstehungskontext, der gemeinsam mit den jeweiligen Eigenleistungen im Folgenden vorgestellt wird. Die Beiträge 1 bis 5 teilen den Kontext des nationalen PISA-Berichtsbandes, der von mir mitherausgegeben wurde.

**Beitrag 1.** Das Einleitungskapitel des nationalen PISA-Berichtsbandes zu PISA 2012 wurde von mir als Erstautorin federführend geplant und zu großen Teilen verfasst. Die Begutachtung erfolgte durch die Herausgeber.

**Beitrag 2.** Das technische Kapitel des PISA-Berichtsbandes wurde vom Erstautor in Zusammenarbeit mit mir konzipiert. Teile des Textes wurden federführend von mir verfasst. Die Begutachtung erfolgte durch die Herausgeber.

**Beitrag 3.** Das Kapitel zur Hauptdomäne in PISA 2012 (mathematische Kompetenz) wurde von mir als Erstautorin entworfen und in weiten Teilen geschrieben (mit Ausnahme des Teils zu den Bildungsstandards und eines Teils der Diskussion). Die im Beitrag berichteten statistischen Analysen wurden von mir durchgeführt und aufbereitet. Die Begutachtung erfolgte durch die Herausgeber.

**Beitrag 4.** Das Kapitel zur Lesekompetenz wurde von der Erstautorin aufgesetzt und hauptverantwortlich geschrieben. Mein Beitrag lag in der Interpretation und Diskussion der Analysen, die ich gemeinsam mit den Ko-Autorinnen verfasste. Darüber hinaus habe ich intensiv an der Überarbeitung des Beitrags nach der Begutachtung durch die Herausgeber mitgewirkt.

**Beitrag 5.** Beim Kapitel zu den schulischen Rahmenbedingungen der Kompetenzentwicklung übernahm ich die Erstautorenschaft. Der Beitrag wurde gemeinsam mit den Ko-Autoren entworfen und

dann von mir federführend umgesetzt (sowohl Datenauswertung als auch Verfassen des Textes). Die Begutachtung erfolgte durch die Herausgeber.

**Beitrag 6.** Der Beitrag ist ein Aufsatz in einem Sonderheft der slowenischen Zeitschrift *Solsko Polje (The School Field)*. Als Erstautorin habe ich den Beitrag anhand der Strukturvorschläge des Ko-Autors konzipiert, die notwendigen Recherchen getätigt und einen Großteil des Textes verfasst. Der Beitrag wurde in einem externen Reviewverfahren begutachtet.

**Beitrag 7.** Dieser Beitrag ist ein Kapitel im Herausgeberband „Internationale Schulleistungsstudien“ (herausgegeben von Gläser-Zikuda, M., Harring, M. & Rohlf, C.). Ich übernahm als Erstautorin die Konzeption des Beitrags sowie das Verfassen eines Textentwurfes. Die Begutachtung erfolgte durch die Herausgeber. Der Band soll Anfang 2016 erscheinen.

**Beitrag 8.** Das Kapitel entstand als Beitrag des im Rahmen des NELSEN-Netzwerkes (*Network of Large-Scale Studies Including Students with Special Educational Needs*) herausgegebenen Sonderbandes zur Inklusion von Schülerinnen und Schülern mit besonderem Förderbedarf. Als Erstautorin des Beitrags habe ich die Idee entwickelt sowie die notwendigen Recherchen getätigt und den Text verfasst. Die Begutachtung erfolgte durch die Herausgeber sowie Autoren anderer Beiträge im selben Band.

**Beitrag 9.** Dieser Beitrag verwendet Daten aus dem ZIB-Begleitforschungsprojekt zu PISA 2012, das eine erweiterte Stichprobe von 49 Förderschulen umfasst. In meiner Rolle als Zweitautorin habe ich nach Rücksprache mit dem Erstautor mehrere Absätze des Textes geschrieben sowie die Analysestrategie und die Gesamtanlage des Beitrags mitentwickelt. Die Begutachtung erfolgte in Form eines externen Reviewprozesses.

**Beitrag 10.** Der Beitrag entstand ebenfalls im Kontext des ZIB-Begleitforschungsprojektes zu PISA 2012, das eine erweiterte Stichprobe von 49 Förderschulen untersucht. An der Entwicklung des Designs sowie an der Erhebung der Daten war ich maßgeblich beteiligt. Teile des Textes habe ich eigenständig verfasst. Die Begutachtung erfolgte in Form eines externen Reviewprozesses.

## Literaturverzeichnis

- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., Schneider, W., Stanat, P., Ostermeier, C., Schiefele, U., Valtin, R. & Ring, K. (2007). *Förderung von Lesekompetenz. Expertise* (Bildungsforschung, Bd. 17). Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–140). Opladen: Leske + Budrich.
- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen* (1. Aufl.). Bielefeld: W. Bertelsmann Verlag.
- Baker, E. T., Wang, M. C. & Walberg, H. J. (1995). The effect of inclusion on learning. *Educational Leadership*, 52 (4), S. 33–35.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W. (Hrsg.). (2002). *Pisa 2000 – die Länder der Bundesrepublik Deutschland im Vergleich. [PISA-E]*. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P. et al. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J. & Köller, O. (1998). Nationale und internationale Schulleistungsstudien: was können sie leisten, wo sind ihre Grenzen? *Pädagogik*, 50 (6), 12–18.
- Bellmann, J. (2006). Bildungsforschung und Bildungspolitik im Zeitalter 'Neuer Steuerung'. *Zeitschrift für Pädagogik*, 52 (4), 487–504.
- Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41 (3), 287–302.  
doi:10.1080/0305764X.2011.607151
- Bieber, T., Martens, K., Niemann, D. & Windzio, M. (2014). Grenzenlose Bildungspolitik? Empirische Evidenz für PISA als weltweites Leitbild für nationale Bildungsreformen. *Zeitschrift für Erziehungswissenschaft*, 17 (S4), 141–166. doi:10.1007/s11618-014-0513-6
- Bischof, L. M., Hochweber, J., Hartig, J. & Klieme, E. (2013). Schulentwicklung im Verlauf eines Jahrzehnts. Erste Ergebnisse des PISA-Schulpanels. In N. Jude & E. Klieme (Hrsg.) *PISA 2009 - Impulse für die Schul- und Unterrichtsforschung. Zeitschrift für Pädagogik*, 59, (S. 172–199) [Themenheft]. Weinheim: Beltz.
- Blanck, J. (2014). *Organisationsformen schulischer Integration und Inklusion: Eine vergleichende Betrachtung der 16 Bundesländer* (WZB Discussion Papers 2014-501). Berlin: WZB.
- Blanck, J., Edelstein, B. & Powell, J. J. W. (2013). Persistente schulische Segregation oder Wandel zur inklusiven Bildung? Die Bedeutung der UN-Behindertenrechtskonvention für Reformprozesse in den deutschen Bundesländern. *Schweizerische Zeitschrift für Soziologie*, 39 (2), 267–292.
- Blömeke, S. & Herzig, B. (2009). Schule als gestaltete und zu gestaltende Institution - ein systematischer Überblick über aktuelle und historische Schultheorien. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtasik & W. Sacher (Hrsg.), *Handbuch Schule. Theorie - Organisation - Entwicklung* (UTB Pädagogik, Bd. 8392, S. 15–28). Bad Heilbrunn: Klinkhardt.
- Blum, W., Drüke-Noe, K., Hartung, R. & Köller, O. (Hrsg.). (2010). *Bildungsstandards Mathematik konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsideen und Fortbildungsmöglichkeiten* (4., vollständig überarbeitete Auflage). Berlin: Cornelsen/Scriptor.

- Bock, R. D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46 (4), 443–459.
- Bock, R. D. & Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, 35 (2), 179–197.
- Bos, W., Bonsen, M. & Gröhlich, C. (Hrsg.). (2009). *KESS 7 - Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7* (HANSE - Hamburger Schriften zur Qualität im Bildungswesen, Bd. 5). Münster: Waxmann.
- Bromme, R., Prenzel, M. & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. Eine Analyse von Anforderungen an die Darstellung, Interpretation und Rezeption empirischer Befunde. *Zeitschrift für Erziehungswissenschaft* 17 (Sonderheft 27), 3–54. Von der Forschung zur evidenzbasierten Entscheidung. Die Darstellung und das öffentliche Verständnis der Bildungsforschung.
- Buchhaas-Birkholz, D. (2009). Die "empirische Wende" in der Bildungspolitik und in der Bildungsforschung. Zum Paradigmenwechsel des BMBF im Bereich der Forschungsförderung. *Erziehungswissenschaft*, 20 (39), 27–33.
- Bundschuh, K. (2006). Rahmenbedingungen und diagnostische Umsetzung zur Feststellung sonderpädagogischen Förderbedarfs in Bayern. In U. Petermann & F. Petermann (Hrsg.), *Diagnostik sonderpädagogischen Förderbedarfs* (Tests und Trends, N.F., Bd. 5, Bd. 5, S. 17–35). Göttingen: Hogrefe.
- Bundschuh, K. & Baier, S. (2009). Sonderschulwesen und sonderpädagogische Einrichtungen. In S. Blömeke (Hrsg.), *Handbuch Schule. Theorie - Organisation - Entwicklung* (UTB Pädagogik, Bd. 8392, S. 288–298). Bad Heilbrunn: Klinkhardt.
- Carnoy, M. (2014). Globalization, educational change, and the national state. In N. P. Stromquist & K. Monkman (Hrsg.), *Globalization and education. Integration and contestation across cultures* (S. 21–38). Lanham, MD: Rowman & Littlefield.
- Carstensen, C. H., Knoll, S., Rost, J. & Prenzel, M. (2004). Technische Grundlagen. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand. (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland: Ergebnisse des zweiten internationalen Vergleichs* (S. 371–388). Münster: Waxmann.
- Cortina, K. S., Baumert, J., Leschinsky, A., Mayer, K. U. & Trommer, L. (2003). *Das Bildungswesen in der Bundesrepublik Deutschland*. Reinbek: Rowohlt.
- Demeris, H., Childs, R. A. & Jordan, A. (2007). The influence of students with special needs included in grade-3 classrooms on the large-scale achievement scores of students without special needs. *Canadian Journal of Education*, 30 (3), 609–627.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. *Zeitschrift für Pädagogik*, (41. Beiheft), 73–92.
- Drechsel, B. & Prenzel, M. (2008). *Aus Vergleichsstudien lernen. Aufbau, Durchführung und Interpretation internationaler Vergleichsstudien* (Schulmanagement-Handbuch, Bd. 126. Jg. 27). München: Oldenbourg.
- Drechsel, B., Prenzel, M. & Seidel, T. (2009). Nationale und internationale Schulleistungstudien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 353–382). Heidelberg: Springer.
- Drechsel, B., Prenzel, M. & Seidel, T. (2014). Nationale und internationale Schulleistungstudien. In E. Wild (Hrsg.), *Pädagogische Psychologie* (2. vollst. überarb. und aktualisierte Aufl., S. 343–368). Berlin: Springer.
- Ehmke, T., Blum, W., Neubrand, M., Jordan, A. & Ulfig, F. (2006). Wie verändert sich die mathematische Kompetenz von der neunten zur zehnten Klassenstufe? In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 63–86). Münster: Waxmann.

- Fay, R. E. (1989). Theory and application of replicate weighting for variance calculations. *Journal of the American Statistical Association*, 91, 343–348.
- Fend, H. (1981). *Theorie der Schule* (2., durchges. Aufl.). München [u.a.]: Urban & Schwarzenberg.
- Fend, H. (2006). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen* (Lehrbuch, 1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fickermann, D. & Maritzen, N. (Hrsg.). (2014). *Grundlagen für eine daten- und theoriegestützte Schulentwicklung. Konzeption und Anspruch des Hamburger Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)*. Münster: Waxmann.
- Fuchs, H.-W. (2003). Auf dem Weg zu einem Weltcurriculum? Zum Grundbildungskonzept von PISA und der Aufgabenzuweisung an die Schule. *Zeitschrift für Pädagogik*, 49 (2), 161–179.
- Gebhardt, M., Heine, J.-H. & Sälzer, C. (2015). Schulische Kompetenzen von Schülerinnen und Schülern ohne sonderpädagogischen Förderbedarf im gemeinsamen Unterricht. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 84 (3), 246. doi:10.2378/vhn2015.art28d
- Gebhardt, M., Heine, J.-H., Zeuch, N. & Förster, N. (2015). Lernverlaufsdagnostik im Mathematikunterricht der zweiten Klasse: Raschanalysen und Empfehlungen zur Adaptation eines Testverfahrens für den Einsatz in inklusiven Klassen. *Empirische Sonderpädagogik*, (3), 206–222.
- Gebhardt, M., Krammer, M., Schwab, S. & Gasteiger-Klicpera, B. (2013). Zusammenarbeit zwischen KlassenlehrerIn und SonderpädagogIn in der Integration: Eine Untersuchung in integrativen Klassen in der Steiermark. *Heilpädagogische Forschung*, 39 (2), 54–61.
- Gebhardt, M., Sälzer, C., Mang, J., Müller, K. & Prenzel, M. (2015). Performance of students with special educational needs in Germany: Findings from Programme for International Student Assessment 2012. *Journal of Cognitive Education and Psychology*, 14 (3), 343–356.
- Gebhardt, M., Sälzer, C. & Tretter, T. (2014). Die gegenwärtige Umsetzung des gemeinsamen Unterrichts in Deutschland. *Heilpädagogische Forschung*, 40 (1), 22–31.
- Gogolin, I. (2010). Chancen und Risiken nach PISA – über Bildungsbeteiligung von Migrant\*innenkindern und Reformvorschläge. In G. Auernheimer (Hrsg.), *Schieflagen im Bildungssystem. Die Benachteiligung der Migrant\*innenkinder* (Interkulturelle Studien, Bd. 16, 4. Aufl., S. 33–50). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gräsel, C. (2011). Was ist Empirische Bildungsforschung? In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (Lehrbuch, 1. Aufl., S. 13–27). Wiesbaden: VS-Verl.
- Gresch, C. & Piezunka, A. (2015). Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich "Lernen") an Regelschulen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 181–220). Wiesbaden: Springer VS.
- Heine, J.-H., Sälzer, C., Borchert, L., Sibberns, H. & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 309–346). Münster: Waxmann.
- Heydrich, J., Weinert, S., Nusser, L., Artelt, C. & Carstensen, C. H. (2013). Including students with special educational needs into large-scale assessments of competencies: challenges and approaches within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online*, 5 (2), 217–240.
- Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 217–244). Münster: Waxmann.
- Klieme, E. (2013). The role of large-scale assessments in research on educational effectiveness and school development. In M. von Davier, E. Gonzalez, I. Kirsch & K. Yamamoto (Hrsg.), *The role of*

- international large-scale assessments: perspectives from technology, economy, and educational research* (S. 115–148). Dordrecht: Springer Netherlands.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards. Expertise*. Bonn: BMBF.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik* 52 (6), 876–903.
- Klieme, E. & Vieluf, S. (2013). Schulische Bildung im internationalen Vergleich. In N. Jude & E. Klieme (Hrsg.) PISA 2009 - Impulse für die Schul- und Unterrichtsforschung. *Zeitschrift für Pädagogik*, 59, 229–246 [Themenheft]. Weinheim: Beltz.
- KMK. (1994). *Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland. Beschluss der Kultusministerkonferenz vom 06.05.1994*, Berlin.
- KMK. (1997). Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland. Konstanzer Beschluss vom 24.10.1997.
- KMK. (2002). PISA 2000 – Zentrale Handlungsfelder. Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen. Beschluss der 299. Kultusministerkonferenz vom 17./18.10.2002.
- KMK. (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (10. Jahrgangsstufe)* (Beschlüsse der Kultusministerkonferenz).
- KMK. (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Wolters Kluwer.
- KMK. (2011). *Inklusive Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen*. Beschluss der Kultusministerkonferenz vom 20.10.2011, Berlin.
- KMK. (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. Beschluss der 350. Kultusministerkonferenz vom 11.06.2015, Berlin.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A. & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66 (2), 165–191. doi:10.1007/s11577-014-0253-x
- Kocaj, A., Kuhl, P., Rjosk, C., Jansen, M., Pant, H. A. & Stanat, P. (2015). Der Zusammenhang zwischen Beschulungsart, Klassenkomposition und schulischen Kompetenzen von Kindern mit sonderpädagogischem Förderbedarf. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 335–370). Wiesbaden: Springer VS.
- Kuhl, P., Stanat, P., Lütje-Klose, B., Gresch, C., Pant, H. A. & Prenzel, M. (Hrsg.). (2015). *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen*. Wiesbaden: Springer VS.
- Kuper, H. (2006). Rückmeldung und Rezeption - zwei Seiten der Verwendung wissenschaftlichen Wissens im Bildungssystem. In H. Kuper & J. Schneewind (Hrsg.), *Rückmeldungen und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungssystem* (S. 7–18). Münster: Waxmann.
- Lange, H. (2008). Vom Messen zum Handeln: "empirische Wende" der Bildungspolitik? *Recht der Jugend und des Bildungswesens*, 56, 7–15.
- Lehmann, R. & Hoffmann, E. (Hrsg.). (2009). *BELLA. Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf "Lernen"*. Münster: Waxmann.
- Lüders, M. & Rauin, U. (2004). *Unterrichts- und Lehr-Lern-Forschung*. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 691–719). Wiesbaden: VS Verlag für Sozialwissenschaften; Imprint: VS Verlag für Sozialwissenschaften.

- Lüke, T. & Ritterfeld, U. (2012). Elterliche Schulzufriedenheit in integrativer und segregativer Beschulung sprachauffälliger Kinder: Ein Vergleich zwischen Förderschule und Gemeinsamen Unterricht. doi:10.17877/DE290R-173
- Martin, M. O. & Mullis, I. V. S. (Hrsg.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L. & Sainsbury, M. (2012). *PIRLS 2011 Assessment Framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y. & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Myklebust, J. O. (2002). Inclusion or exclusion? Transition among special needs students in upper secondary education in Norway. *European Journal of Special Needs Education*, 17 (3), 251–263.
- Neugebauer, M. (2010). Bildungsungleichheit und Grundschulempfehlung beim Übergang auf das Gymnasium: Eine Dekomposition primärer und sekundärer Herkunftseffekte. *Zeitschrift für Soziologie* 39, (3), 202–214. doi:10.2307/23773069
- Nusser, L., Carstensen, C. H. & Artelt, C. (2015). Befragungen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen: Ergebnisse zur Messinvarianz. *Empirische Sonderpädagogik*, 7 (2), 99–116.
- OECD. (2013a). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris: OECD Publishing.
- OECD. (2013b). *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices. Volume IV*. Paris: OECD Publishing.
- OECD. (2014a). *PISA 2012 Results: What students know and can do. Student performance in mathematics, reading and science* (Revised edition). Paris: OECD Publishing.
- OECD. (2014b). *PISA 2012 Technical Report*. Paris: OECD Publishing.
- Ostermeier, C., Prenzel, M. & Duit, R. (2010). Improving Science and Mathematics instruction - The SINUS-Project as an example for reform as teacher professional development. *International Journal of Science Education*, 32 (3), 303–327.
- Pestalozza, C. (Hrsg.). (2014). *Verfassungen der deutschen Bundesländer* (Beck'sche Textausgaben, 10. Aufl.). Beck'sche Textausgaben. München: Beck.
- Petermann, U. & Petermann, F. (Hrsg.). (2006). *Diagnostik sonderpädagogischen Förderbedarfs* (Tests und Trends, N.F., Bd. 5). Göttingen: Hogrefe.
- Pfahl, L. (2012). Bildung, Behinderung und Agency. Eine wissenssoziologische Untersuchung der Folgen schulischer Segregation und Inklusion. In R. Becker & H. Solga (Hrsg.), *Soziologische Bildungsforschung* (Kölner Zeitschrift für Soziologie und Sozialpsychologie : Sonderheft, Bd. 52, S. 415–436). Wiesbaden: Springer Fachmedien Wiesbaden; Imprint: Springer VS.
- Pohl, S. & Carstensen, C. H. (2013). Scaling of competence tests in the National Educational Panel Study -- Many questions, some answers, and further challenges. *Journal for Educational Research Online*, 5 (2), 189–216.
- Prenzel, M. (2005). Zur Situation der empirischen Bildungsforschung. In H. Mandl & B. Kopp (Hrsg.), *Impulse für die Bildungsforschung: Stand und Perspektiven Dokumentation eines Expertengesprächs* (S. 7–21). De Gruyter.
- Prenzel, M. (2009). Von der Unterrichtsforschung zur Exzellenz in der Lehrerbildung. *Beiträge zur Lehrerbildung*, 27 (3), 327–345.

- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. et al. (Hrsg.). (2008). *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich*. Münster: Waxmann.
- Prenzel, M. & Baumert, J. (Hrsg.). (2008) Vertiefende Analysen zu PISA 2006 [Themenheft] 10.2008. Wiesbaden: VS Verl. für Sozialwiss.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. & Schiefele, U. (2006). *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster: Waxmann.
- Prenzel, M., Friedrich, A. & Stadler, M. (Hrsg.). (2009). *Von SINUS lernen. Wie Unterrichtsentwicklung gelingt* (Sinus-Transfer, 1. Aufl.). Seelze: Kallmeyer.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (Hrsg.). (2013). *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Rammstedt, B. & Spinath, F. M. (2013). Öffentliche Datensätze und ihr Mehrwert für die psychologische Forschung. *Psychologische Rundschau*, 64 (2), 101–102. doi:10.1026/0033-3042/a000156
- Reise, S. P. & Revicki, D. A. (Hrsg.). (2014). *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment* (Multivariate Application Series, 1. Aufl.). London: Routledge.
- Roeder, P. M. (2003). TIMSS und PISA- Chancen eines neuen Anfangs in Bildungspolitik, -Planung, -Verwaltung und Unterricht. Endlich ein Schock mit Folgen? *Zeitschrift für Pädagogik* 49 (2), 180–197.
- Rolff, H.-G. (2000). *Manual Schulentwicklung. Handlungskonzept zur pädagogischen Schulentwicklungsberatung*. Weinheim: Beltz.
- Rust, K. F. & Rao, J. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5 (3), 283–310. doi:10.1177/096228029600500305
- Rutkowski, D. J. & Prusinski, E. L. (2011). The limits and possibilities of international large-scale assessments, Center for Evaluation & Education Policy. Education Policy Briefs: Vol. 9, No. 2.
- Sälzer, C. (2015). *Studienbuch Schulleistungsstudien. Das Rasch-Modell in der Praxis*. Berlin: Springer.
- Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 129–152). Wiesbaden: Springer VS.
- Sälzer, C. & Prenzel, M. (2013). PISA 2012 – eine Einführung in die aktuelle Studie. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 11–46). Münster: Waxmann.
- Sälzer, C. & Prenzel, M. (2014). Looking back at five rounds of PISA: impacts on teaching and learning mathematics in Germany. *Solsko Polje (The School Field)*, 25 (5-6), 53–72.
- Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 155–187). Münster: Waxmann.
- Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Mathematische Kompetenz im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 47–98). Münster: Waxmann.

- Sälzer, C., Schiepe-Tiska, A. & Prenzel, M. (in Druck). Internationale Schulleistungsstudien. In M. Gläser-Zikuda, M. Harring & C. Rohlf's (Hrsg.), *Handbuch Schulpädagogik*. Münster: Waxmann.
- Schiepe-Tiska, A., Schmidtner, S. & Prenzel, M. (2014). *Naturwissenschaftsleistungen 15-Jähriger: Ergebnisse aus PISA 2012* (Bd. 4).
- Schuck, K. D. (2014). Individualisierung und Standardisierung der inklusiven Schule - ein unauflösbarer Widerspruch? *Die Deutsche Schule*, 106 (2), 162–174.
- Seidel, T. & Prenzel, M. (2008). Large scale assessment. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 279–304). Göttingen: Hogrefe & Huber.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31, 15–21.
- Stanat, P. & Pant, H. A. (2014). Hat PISA die Schulen besser gemacht? In B. Spinath (Hrsg.), *Empirische Bildungsforschung* (S. 21–37). Berlin: Springer.
- Straus, M. (2014). The timeless questions about educational quality. *Solsko Polje (The School Field)* XXV, (5-6), 7–10.
- Stufflebeam, D. L. (1983). The CIPP Model for Program Evaluation. In G. F. Madaus, M. S. Scriven & D. L. Stufflebeam (Hrsg.), *Evaluation Models. Viewpoints on Educational and Human Services Evaluation* (Evaluation in Education and Human Services, Bd. 6, Bd. 6, S. 117–141). Dordrecht: Springer Netherlands.
- Tenorth, H.-E. (2004). Bildungsstandards und Kerncurriculum. Systematischer Kontext, bildungstheoretische Probleme. *Zeitschrift für Pädagogik*, 50 (5), 650–661.
- Terhart, E. (2001). Bildungsforschung, Bildungsadministration, Bildungswirklichkeit: eine systematische Annäherung. In K.-J. Tillmann & W. Vollstädt (Hrsg.), *Politikberatung durch Bildungsforschung. Das Beispiel: Schulentwicklung in Hamburg* (Reihe Schule und Gesellschaft, Bd. 27, Bd. 27, S. 17–32). Opladen: Leske + Budrich.
- Tillmann, K. J., Dederich, K., Kneuper, D., Kuhlmann, C. & Nessel, I. (2008). *PISA als bildungspolitisches Ereignis: Fallstudien in vier Bundesländern*: VS Verlag für Sozialwissenschaften.
- Tillmann, K.-J. & Vollstädt, W. (Hrsg.). (2001). *Politikberatung durch Bildungsforschung. Das Beispiel: Schulentwicklung in Hamburg* (Reihe Schule und Gesellschaft, Bd. 27). Opladen: Leske + Budrich.
- United Nations Office of Legal Affairs. (2006). *Convention on the rights of persons with disabilities*. U.N.Doc.: A/61/611. Verfügbar unter: <http://www.un.org/disabilities/convention/conventionfull.shtml>. Zugegriffen 23.04.2014.
- van Ackeren, I., Klemm, K. & Kühn, S. M. (Hrsg.). (2015). *Entstehung, Struktur und Steuerung des deutschen Schulsystems. Eine Einführung* (3., überarb. u. aktualisierte Aufl.). Wiesbaden: Springer VS.
- Wocken, H. & Gröhlich, C. (2009). Kompetenzen von Schülerinnen und Schülern an Hamburger Förderschulen. In W. Bos, M. Bonsen & C. Gröhlich (Hrsg.), *KESS 7 - Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7* (HANSE - Hamburger Schriften zur Qualität im Bildungswesen, Bd. 5, S. 133–142). Münster: Waxmann.
- Yamamoto, K. & Mazzeo, J. (1992). Chapter 4: Item Response Theory Scale Linking in NAEP. *Journal of Educational and Behavioral Statistics*, 17 (2), 155–173.  
doi:10.3102/10769986017002155

# Lebenslauf

Dr. Christine C. Sälzer

## Forschungsschwerpunkte

- Schulleistungsstudien (Large-Scale Assessments)
- Schuleffektivität und Schulentwicklung
- Schulabsentismus
- Messung von Kompetenz
- Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

## Abschlüsse

|      |  |
|------|--|
| 2013 | Zusatzqualifikation Multiplikatorenprogramm in der Hochschullehre (Institut ProLehre der TU München) |
| 2010 | Promotion (Dr. phil.) Erziehungswissenschaften, Universität Freiburg i. Ue.                          |
| 2006 | Lizentiat (lic. phil.) Erziehungswissenschaften, Universität Freiburg i. Ue.                         |
| 2000 | Abitur (Albert-Schweizer-Gymnasium Crailsheim)   |

## Praktika und berufliche Tätigkeit

|             |  |
|-------------|--|
| Aktuell     | Seit 2010 Wissenschaftliche Angestellte an der TU München, Zentrum für Internationale Vergleichsstudien. Nationales Projektmanagement des <i>Programme for International Student Assessment</i> (PISA); Führung von 8 wissenschaftlichen Mitarbeitenden und 6 studentischen Hilfskräften, freiwillige Ausbringung von Lehrveranstaltungen für Bachelor- und Masterstudiengänge an der TU München sowie der LMU München |
| 06/2014     | Forschungsaufenthalt an der School of Education der Stanford University zum Thema fachspezifischer Schulabsentismus und mathematische Kompetenz in PISA 2012   |
| 2013 – 2017 | Part-Time Consultant (nebenberuflich, 52 Arbeitstage) der Gesellschaft für Internationale Zusammenarbeit (GIZ) im Bereich <i>Capacity Building and Education</i> in Prishtina, Kosovo  |
| 2012        | Entwicklung von Skalen zur Erfassung von Interesse und Motivation an naturwissenschaftlichen Fragestellungen für PISA 2015 im Auftrag der OECD   |
| 2009 – 2010 | Wissenschaftliche Mitarbeiterin an der Universität Tübingen, Abteilung für Empirische Bildungsforschung und Pädagogische Psychologie (Prof. Dr. Ulrich Trautwein und Prof. Dr. Oliver Lüdtke); Projektleiterin der Mehrkohorten-Längsschnittstudie TRAIN (3 wissenschaftliche Mitarbeitende und 5 studentische Hilfskräfte, Lehrdeputat: 3 SWS)  |

|             |  |
|-------------|--|
| 2008 – 2009 | <i>Visiting Scholar</i> an der School of Education der Stanford University, Palo Alto, Kalifornien (USA); Arbeit an der Dissertation zum Thema schulische Erklärungsfaktoren für Schulabsentismus                        |
| 2006 – 2008 | Wissenschaftliche Mitarbeiterin und Doktorandin am Departement für Erziehungswissenschaften der Universität Freiburg i. Ue.; Projekte: Schulabsentismus in der Schweiz, Schulabbruch in der Schweiz (Lehrdeputat: 2 SWS) |
| 2004 – 2006 | Unterassistentin am Departement für Erziehungswissenschaften der Universität Freiburg i. Ue.   |
| 2003 – 2004 | Studentische Hilfskraft am Departement für Erziehungswissenschaften der Universität Freiburg i. Ue.  |
| 2003        | Praxissemester in der Beratungsabteilung <i>Siemens Procurement and Logistics Services, Supply Chain Consulting</i> in München; eigenes Projekt zur standortweiten Reduktion von Reisekosten                             |
| 2002 – 2003 | Studentische Hilfskraft am Institut für Pädagogische Psychologie der TU Braunschweig   |
| 2000 – 2001 | Freiwilliges Soziales Jahr in der stationären Jugendhilfe, Bereich Verhaltensauffälligkeiten (Evangelische Jugendhilfe Friedenshort e. V., Öhringen)   |

### **Universitäre Lehre**

|           |   |
|-----------|---|
| Seit 2013 | TUM School of Education (Scientific Writing, Masterstudiengang, 2 SWS)  |
| 2012      | Ludwig-Maximilians-Universität München (Educational Systems and Achievement, Masterstudiengang, 2 SWS)  |
| 2011      | Fachhochschule Nordwestschweiz (Large-Scale Assessments und Schulabsentismus, Lehramtsstudiengang, 1 SWS)   |
| 2009-2010 | Abteilung Empirische Bildungsforschung und Pädagogische Psychologie am Institut für Erziehungswissenschaften der Universität Tübingen (Quantitative Datenanalyse, Bachelorstudiengang, 2 SWS; Schulabsentismus, Lehramtsstudiengang für Gymnasien, 2 SWS) |
| 2006-2008 | Departement für Erziehungswissenschaften, Universität Freiburg i. Ue. (Geschichte der Pädagogik, Forschungsworkshop sowie Wissenschaftliches Arbeiten, Bachelorstudiengang, jeweils 1 SWS)  |

### **Aktuelle Projekte**

- PISA (*Programme for International Student Assessment*) 2012, 2015 und 2018
- PISA Plus (Messwiederholung ein Jahr nach PISA 2012, drittmittelgefördert im Rahmen des Forschungsbudgets des Zentrums für Internationale Vergleichsstudien; ca. 550.000€)

- Zusatzerhebung zu PISA 2012: Erweiterte Stichprobe mit 49 Sonder- und Förderschulen (drittmittelgefördert im Rahmen des Forschungsbudgets des Zentrums für Internationale Vergleichsstudien; ca. 100.000€)
- Evaluation des Berliner Projekts Schuldistanz 2014 (drittmittelgefördert von der Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin, 10.000€)
- Finale Antragsphase: Erasmus+ Programme der Europäischen Union, European Policy Experimentation. Kooperation zur Erweiterung der PISA-2015-Erhebung (ca. Herbst 2015) in Zusammenarbeit mit den PISA National Centres in Slowenien, Schweden und Polen

## **Auszeichnungen und Stipendien**

|             |  |
|-------------|--|
| 2015        | 2. Listenplatz im Berufungsverfahren auf eine assoziierte Professur für Empirische Schul- und Unterrichtsforschung am Departement für Erziehungswissenschaften der Universität Fribourg (Schweiz)                |
| 2010-2013   | Reisestipendien des DAAD zur Teilnahme an Konferenzen in Belgien, den Niederlanden, Großbritannien und den USA   |
| 2008 – 2009 | Stipendium für angehende Forschende des Schweizerischen Nationalfonds zur Förderung der Wissenschaft; Gastaufenthalt an der School of Education der Stanford University, Palo Alto, Kalifornien (ca. 50.000 SFR) |
| 2008        | Förderung des Schweizerischen Nationalfonds zur Teilnahme an der Essex Summer School (3.000 SFR)   |
| 2007        | Stipendium der British Educational Research Association zur Teilnahme an der BERA Annual Conference am Institute of Education, London (UK)   |
| 2007        | Förderung des Schweizerischen Nationalfonds zur Teilnahme an der Essex Summer School (3.000 SFR)   |
| 2003 – 2004 | Mobilitätsstipendium des europäischen Erasmus-Austauschprogrammes für den Studienaufenthalt in Freiburg i. Ue. (3.600€)  |

## **Mitgliedschaften (in alphabetischer Reihenfolge)**

American Educational Research Association (AERA), Deutsche Gesellschaft für Erziehungswissenschaft (DGfE), Deutscher Hochschulverband (DHV), European Association for Research on Learning and Instruction (EARLI), Gesellschaft für Empirische Bildungsforschung (GEBF), Schweizerische Gesellschaft für Bildungsforschung (SGBF)

## **Gutachtertätigkeiten (in alphabetischer Reihenfolge)**

AERA Annual Conference, Bundesministerium für Bildung und Forschung (Begutachtung von Projektanträgen), Gesellschaft für Empirische Bildungsforschung (GEBF), Journal of Cognitive Education and Psychology, Journal of Educational Psychology, Learning and Individual Differences, Learning and Instruction, OECD (Organisation for Economic Co-operation and Development): Mitglied

des Technical Review Panels für PISA 2018, Technology, Knowledge and Learning, Unterrichtswissenschaft, Zeitschrift für Pädagogische Psychologie

### **Akademische Selbstverwaltung**

- |             |   |
|-------------|---|
| Seit 2012   | Gründungsmitglied des NELSEN-Netzwerks (NeEtnetwork of Large-Scale Studies Including Students with Special Educational Needs)                             |
| Seit 2011   | Auswahlgespräche für Studienbewerber an der TUM School of Education (Lehramt Naturwissenschaftliche Bildung, Master of Research on Teaching and Learning) |
| 2012 – 2013 | Beteiligung an der Organisation der Konferenz der European Association for Research on Teaching and Learning (EARLI) 2013 in München                      |
| 2011        | Mitarbeit an der Konzeption des Masterstudiengangs Research on Teaching and Learning  |
| 2007 – 2008 | Studienberaterin am Departement für Erziehungswissenschaften der Universität Freiburg i. Ue.  |
| 2007-2008   | Gewählte Vertreterin des Mittelbaus im Fakultätsrat der Philosophischen Fakultät der Universität Freiburg i. Ue.  |

# Publikationen und Vorträge

## Monographien und Herausgeberschaften (\*peer-reviewed)

Sälzer, C. (in Druck). *Studienbuch Schulleistungsstudien. Das Rasch-Modell in der Praxis*. Berlin: Springer. DOI 0.1007/978-3-662-45765-8

\*Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

Sälzer, C. (2010). *Schule und Absentismus. Individuelle und schulische Faktoren für jugendliches Schwänzverhalten*. Wiesbaden: VS Verlag für Sozialwissenschaften (Dissertation).

Stamm, M., Ruckdäschel, C., Templer, F. & Niederhauser, M. (2009). *Schulabsentismus. Ein Phänomen, seine Bedingungen und Folgen*. Wiesbaden: VS Verlag für Sozialwissenschaften.

## Artikel und Buchbeiträge (\*peer-reviewed)

\*Sälzer, C. & Mang, J. (revise & resubmit). Re-thinking truancy in the context of Item Response Theory (IRT): A scaling-approach for a behavioral measure in a student questionnaire. *Frontiers in Psychology*.

\*Sälzer, C. & Lenski, A. E. (revise & resubmit). School attendance policies: How students differentially perceive their school's reactions to truant behavior. *Journal of Education for Students Placed at Risk*.

\*Sälzer, C. & Heine, J.-H. (revise & resubmit). Kompetenzerwerb im Unterricht: Zum Zusammenhang von gezieltem Schwänzen von Mathematikstunden und Schülerleistungen im Mathematiktest in PISA 2012 in unterschiedlichen Sekundarschularten. *Psychologie in Erziehung und Unterricht*.

\*Müller, K., Prenzel, M., Sälzer, C., Gebhardt, M. & Mang, J. (revise & resubmit). Deskriptive Befunde zum Leistungsniveau und zur Herkunft der Sonder- und Förderschüler in Deutschland. *Unterrichtswissenschaft*.

\*Sälzer, C., Schiepe-Tiska, A. & Prenzel, M. (revise & resubmit). Internationale Schulleistungsstudien. In: M. Gläser-Zikuda, M. Harring & C. Rohlf's (Hrsg.), *Handbuch Schulpädagogik*. Münster: Waxmann.

\*Sälzer, C. & Heine, J.-H. (in press). Students' skipping behavior on truancy items and (school) subjects and its relation to test performance in PISA 2012. *International Journal of Educational Development*. DOI 10.1016/j.ijedudev.2015.10.009

\*Piezunka, A., Gresch, C., Sälzer, C. & Kroth, A. (in Druck). Identifizierung und Erfassung von Schülerinnen nach Vorgaben der UN-BRK im Rahmen von bundesweiten Erhebungen: Herausforderungen und Perspektiven. *Zeitschrift für Pädagogik, Beiheft Schulische Inklusion*.

\*Gebhardt, M., Sälzer, C., Mang, J., Müller, K., & Prenzel, M. (2015). Performance of Students with Special Educational Needs in Germany. Findings from PISA 2012. *Journal of Cognitive Education and Psychology*, 14 (3), 343-356.

\*Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In: P. Kuhl, P. Stanat, B. Lütje-Klose, C.

- Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Forschungsstand, Herausforderungen und Perspektiven* (S. 129-152). Wiesbaden: Springer.
- \*Gebhardt, M., Heine, J.-H. & Sälzer, C. (2015). Schulische Kompetenzen von Schülerinnen und Schülern ohne sonderpädagogischen Förderbedarf im gemeinsamen Unterricht. *Vierteljahresschrift für Heilpädagogik*, 84 (3), 246-258.
- Sälzer, C. (2015). Schulschwänzen bei Neuntklässlern: Warum, wie oft und wie? *Schule im Blickpunkt*, 48, 8-10.
- \*Zehner, F., Sälzer, C. & Goldhammer, F. (2015). Automatic Coding of Short Text Responses via Clustering in Educational Assessment. *Educational and Psychological Measurement*, published online before print. DOI 10.1177/0013164415590022.
- \*Van den Ham, A.-K., Nissen, A., Ehmke, T., Sälzer, C. & Roppelt, A. (2014). Mathematische Kompetenz in PISA, IQB-Ländervergleich und NEPS – drei Studien, gleiches Konstrukt?. *Unterrichtswissenschaft*, 42 (4), 321-341.
- \*Sälzer, C. & Prenzel, M. (2014). Looking Back at Five Rounds of PISA: Impacts on Teaching and Learning Mathematics in Germany. *Solsko Polje (The School Field)*, 25 (5-6), 53-72.
- \*Gebhardt, M., Schwab, S., Krammer, Gasteiger-Klicpera, B. & Sälzer, C. (2014). Erfassung von individualisiertem Unterricht in der Sekundarstufe I. Eine Quantitative Überprüfung der Skala „Individualisierter Unterricht“ in zwei Schuluntersuchungen in der Steiermark. *Zeitschrift für Bildungsforschung*, online first, DOI 10.1007/s35834-014-0095-7.
- Gebhardt, M., Sälzer, C. & Tretter, M. (2014). Die gegenwärtige Umsetzung des gemeinsamen Unterrichts in Deutschland. *Heilpädagogische Forschung*, 40 (1), 22-31.
- Sälzer, C. & Prenzel, M. (2014). Financial Literacy im Rahmen der PISA-Studie. In: Retzmann, Th. (Hrsg.) *Ökonomische Allgemeinbildung in der Sekundarstufe I und Primarstufe. Konzepte, Analysen, Studien und empirische Befunde*, (S. 15-31). Schwalbach/Ts.: Wochenschau Verlag.
- Sälzer, C., als Mitglied der Autorengruppe Bildungsberichterstattung (2014): *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- \*Heine, J.-H., Sälzer, C., Borchert, L., Sibberns, H. & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. (S. 309-346). Münster: Waxmann.
- \*Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. (S. 217-244).Münster: Waxmann.
- \*Sälzer, C. & Prenzel, M. (2013). PISA 2012 - eine Einführung in die aktuelle Studie. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. (S. 11-46). Münster: Waxmann.
- \*Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. (S.155-188). Münster: Waxmann.

- \*Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenz im internationalen Vergleich. In: M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. (S. 47-98).Münster: Waxmann.
- Sälzer, C. & Prenzel, M. (2013). Financial Literacy als Teil der Grundbildung in PISA. *Unterricht Wirtschaft + Politik*, 3, S. 55-57.
- \*Sälzer, C., Trautwein, U., Lüdtke, O. & Stamm, M. (2012): Predicting Adolescent Truancy: The Importance of Distinguishing Between Different Aspects of Instructional Quality. *Learning and Instruction*, 22, 311-319, DOI 10.1016/j.learninstruc.2011.12.001.
- Moroni, S., Düggeli, A. & Sälzer, C. (2010). "Ich bin dann mal weg..." -- Fragen zur Definition, zum Ausmaß und zur unterrichtlichen Relevanz von schuldistanzierendem Verhalten bei Jugendlichen. *Beiträge zur Lehrerbildung*, 28, 405-415.
- \*Sälzer, C. (2009). Schulabsentismus und die Bedeutung von Schule und Familie. *Schweizerische Zeitschrift für Bildungswissenschaften*, 03/09, 625-640.
- \*Stamm, M., Ruckdäschel, C. & Templer, F. (2009). Facetten des Schulschwänzens: empirische Befunde zu schulabsenten Verhaltensformen Jugendlicher. *Diskurs Kindheits- und Jugendforschung*, 01/09, 107-122.
- Stamm, M., Templer, F. & Ruckdäschel, C. (2006): Schulschwänzen – Ein unterschätztes Phänomen? *Universitas, Das Magazin der Universität Freiburg*, Schweiz, Vol. 65/2, 36-37.

### **Vorträge und Kongressbeiträge (\*eingeladen)**

- \*Sälzer, C. (2015). *Schulabsentismus und Schulverweigerung in Deutschland. Empirische Befunde und Interventionen*. Keynote auf der Abschlusskonferenz des Comenius-Projekts "Schulerfolg sichern in Leipzig und Riga" in Riga.
- Bauer, J., Sälzer, C. & Prenzel, M. (2015). *Teachers' attitudes toward evidence-based practice: The role of access to research resources and self-concept about understanding research*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in Chicago, Illinois.
- Sälzer, C. & Heine, J.-H. (2015). *Subject-specific truancy: Dimensionality of the truancy-scale and associations with subject-specific achievement*. Poster presented at the Annual Meeting of the American Educational Research Association (AERA) in Chicago, Illinois.
- Zehner, F., Sälzer, C. & Goldhammer, F. (2015). *How automatic coding of short text responses is able to enhance assessment*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in Chicago, Illinois.
- Sälzer, C., Seidel, T. & Bayer, S. (2015). *Der Zusammenhang zwischen Kompetenzentwicklung und dem Gefühl der Zugehörigkeit zur Schule in unterschiedlichen Sekundarschulformen*. Vortrag gehalten auf der 3. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) in Bochum.
- Reiss, K., Heinze, A., Sälzer, C., Nagy, G., Pant, H. A., Kiemer, K. & Lehner, M. (2015). *Veränderung der mathematischen Kompetenz von der neunten zur zehnten Klassenstufe*. Vortrag gehalten auf der 3. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) in Bochum.

- Hertel, S., Jude, N. & Sälzer, C. (2015). *Die Zusammenarbeit von Lehrkräften und Eltern im Rahmen von Elterngesprächen – Befunde aus PISA 2012*. Vortrag gehalten auf der 3. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) in Bochum.
- Zehner, F., Sälzer, C. & Goldhammer, F. (2015). *Wie automatisches Kodieren offener Kurztextantworten zur Verbesserung von Assessments beitragen kann*. Vortrag gehalten auf der 3. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) in Bochum.
- Van den Ham, A.-K., Ehmke, T., Schroeders, U. & Sälzer, C. (2014). *Validität der NEPS-Mathematiktests für die neunte Klasse*. Vortrag gehalten auf dem 49. Kongress der Deutschen Gesellschaft für Psychologie (DGPs) in Bochum.
- Bauer, J., Sälzer, C. & Prenzel, M. (2014). *Einstellungen von Lehrkräften gegenüber pädagogischer Forschung und evidenzbasiertem Unterricht*. Vortrag gehalten auf dem 49. Kongress der Deutschen Gesellschaft für Psychologie (DGPs) in Bochum.
- \*Sälzer, C. (2014): *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Vortrag im Rahmen der Fachtagung der GPF in Frankfurt/Main.
- \*Sälzer, C. (2014). *Schulabsentismus und Kompetenzerwerb*. Keynote gehalten auf der Jahrestagung des Adipositasnetzwerks Hessen e. V. in der Comeniuschule für Kranke in Bad Orb.
- Sälzer, C. & Heine, J.-H. (2014). *Missing mathematics classes: Does truancy affect students' mathematical literacy?* Vortrag gehalten im Rahmen des 4<sup>th</sup> Biennial Meeting of the EARLI SIG 18 (Educational Effectiveness) in Southampton, Großbritannien.
- Sälzer, C., Heine, J.-H. & Prenzel, M. (2014). *Ohne regelmäßigen Mathematikunterricht keine Kompetenz? Der Zusammenhang von mathematischer Kompetenz und regelmäßigem Unterrichtsbesuch*. Vortrag gehalten auf dem Kongress der Schweizerischen Gesellschaft für Bildungsforschung 2014 in Luzern.
- \*Sälzer, C. & Jude, N. (2014). *PISA 2012: Ergebnisse des 5. Internationalen Vergleichs und Perspektiven für Deutschland*. Vortrag im Rahmen des Symposiums "PISA 2012 – Aktuelle Perspektiven zur fünften Erhebungsrunde" auf der zweiten GEBF-Tagung 2014 in Frankfurt/Main.
- Zehner, F., Sälzer, C. & Goldhammer, F. (2014). *Automatische Auswertung offener, natürlichsprachlicher Antworten in PISA-Testaufgaben*. Poster präsentiert auf der zweiten GEBF-Tagung 2014 in Frankfurt/Main.
- \*Sälzer, C. & Prenzel, M. (2013). *Financial Literacy im Rahmen von PISA*. Keynote gehalten bei der Jahrestagung der Deutschen Gesellschaft für ökonomische Bildung (DeGöB) in Nürnberg.
- Lenski, A. & Sälzer, C. (2013). *Why do some students play truant and others don't? The role of domain-specific self-concept*. Vortrag auf dem AERA Annual Meeting 2013 in San Francisco, USA.
- Müller, K., Sälzer, C., Mang, J., Gebhardt, M. & Prenzel, M. (2013). *Aufmerksamkeitsbezogenes Aufgabenverhalten bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf*. Vortrag im Rahmen der ersten Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) in Kiel.
- Sälzer C. (2013). *"I just needed some sleep". Using student self-reports in a large-scale assessment to find major reasons for playing truant*. Vortrag im Rahmen des Symposiums "At-risk students in Large-Scale Assessments" auf der EARLI Biennial Conference in München.

- Sälzer, C. & Mang, J. (2013). *Re-thinking student behavior: Truancy from an IRT perspective. Evaluated and compared through a Swiss Large-Scale Assessment on truancy and PISA 2012 in Germany.* Vortrag auf dem AERA Annual Meeting 2013 in San Francisco, USA.
- Sälzer, C., Lenski, A. & Hofherr, S. (2013). *Schulen gegen Absentismus. Unterschiedliche Strategien und ihre Wirksamkeit.* Vortrag im Rahmen der ersten Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF) in Kiel.
- Sälzer, C., Lenski, A. & Hofherr, S. (2013). *Strategies to fight truancy. How students are prevented from putting themselves at risk.* Vortrag im Rahmen des Symposiums "At-risk students in Large-Scale Assessments" auf der EARLI Biennial Conference in München.
- \*Sälzer, C. (2012). *„Ich musste mal wieder ausschlafen.“ Strategien von Schulen zum Umgang mit Schulabsentismus.* Referat im Rahmen der Lehrerbildung der Fachhochschule Nordwestschweiz, Basel (Schweiz).
- \*Sälzer, C. (2012). *„Ohne mich!“ Schulschwänzen als Forschungsgegenstand.* Vortrag im Rahmen des BiKs-Kolloquiums an der Universität Bamberg (Prof. Cordula Artelt).
- \*Sälzer, C. (2012). *Von Europa lernen: Schulvermeidung – Konzepte und Erfahrungen.* Abschlussreferat auf der Fachtagung „Schulmüde Jugendliche – Was tun?“ des Kooperationsverbundes Jugendsozialarbeit in Fulda.
- \*Sälzer, C., Lenski, A., & Hofherr, S. (2012). *Ohne mich! Konzepte zum Umgang mit Schulverweigerung und ihre Wirksamkeit.* Auftaktvortrag auf der Tagung „Reintegration von Schüler/innen bei Schulversäumnissen in Zusammenarbeit mit dem Projekt 2. Chance“ in München.
- Sälzer, C. (2012). *Can school effectiveness be measured by student attendance rates?* Paper presented at the EARLI SIG 18 Educational Effectiveness Conference, Zurich (Switzerland).
- Sälzer, C., & Mang, J. (2012). *Re-thinking student behavior: truancy in an IRT context.* Paper presented at the AEA-Europe conference in Berlin.
- Sälzer, C., Trautwein, U., Lüdtke, O., & Stamm, M. (2012). *Do truants feel overtaxed at school?* Poster presented at the 93rd annual meeting of the American Educational Research Association (AERA), Vancouver (Canada).
- Sälzer, C. (2011). *Adolescent Truancy – Individual Judgments and Classroom Factors.* Vortrag gehalten auf dem EARLI Biennial Meeting 2011 in Exeter, Grossbritannien.
- Sälzer, C. (2011). *Predicting Adolescent Truancy. The Role of Individual and Classroom Characteristics.* Vortrag gehalten auf dem AERA Annual Meeting 2011 in New Orleans, USA.
- Sälzer, C. (2010). *Truancy – A useful measure of educational effectiveness?* Vortrag gehalten auf dem Second Meeting of the EARLI SIG 18 (Educational Effectiveness) in Leuven, Belgien.
- Sälzer, C. (2010). *Truancy in Switzerland. How Schools Matter. Presentation focusing school climate as a predictor for truant behavior.* AERA Annual Meeting 2010 in Denver, Colorado, USA.
- Sälzer, C. (2009). *Truancy – A Challenge for Education. Results from a Multilevel Analysis.* Vortrag gehalten auf dem EARLI Biennial Meeting 2009 in Amsterdam, Niederlande.
- Ruckdäschel, C. (2008). *Die Rolle schulischer Faktoren bei Schulabsentismus. Vertiefung einer Grundlagenstudie.* Vortrag gehalten auf dem Kongress der DGfE in Dresden.

Ruckdäschel, C. (2007). *The Phenomenon of Truancy and its Consequences*. Vortrag gehalten im Rahmen der BERA (British Educational Research Association) Annual Conference am London Institute of Education.

## **Organisierte Symposien**

Sälzer, C. & Jude, N (2014). "PISA 2012 – Aktuelle Perspektiven zur fünften Erhebungsrunde". Symposium auf der zweiten GEBF-Tagung 2014 in Frankfurt/Main.

### *Beiträge im Rahmen des Symposiums:*

Sälzer, C. & Jude, N. (2014). *PISA 2012: Ergebnisse des 5. Internationalen Vergleichs und Perspektiven für Deutschland*.

Kröhne, U., Goldhammer, F. & Hahnel, C. (2014). *Lassen sich Unterschiede zwischen paper- und computerbasiert gemessener PISA-Lesekompetenz durch Log- und Prozessdaten erklären?*

Schiepe-Tiska, A., Heine, J.-H., Obersteiner, A. & Reiss, K. (2014). *Merkmale des Mathematikunterrichts in Deutschland und ihr Zusammenhang mit kognitiven sowie motivational-affektiven Lernergebnissen in PISA 2012*.

Müller, K., Sälzer, C., Mang, J. & Prenzel, M. (2014). *Kompetenzen von Schülerinnen und Schülern mit besonderem Förderbedarf. Ergebnisse aus dem PISA 2012 Förderschul-Oversample*.

Müller, K. & Sälzer, C. (2013). „At-Risk Students in Large Scale Assessments“. Symposium auf dem EARLI Biennial Meeting 2013 in München.

### *Beiträge im Rahmen des Symposiums:*

Dueggeli, A., Moroni, S. & Kandzia, W. (2013). *Risky neighborhoods: Does a family's address affect students' assignment to different school tracks?*

Ehmke, T., Drechsel, B. & Carstensen, C. (2013). *Explaining the learning progress of retained students and low-achieving students in mathematics*.

Sälzer, C., Lenski, An E. & Hofherr, S. (2013). *Strategies to fight truancy: how students are prevented from putting themselves at risk*.

Gebhardt, M. & Müller, K. (2013). *Assessment options for students with special educational needs in large scale assessments*.

## **Fachpublikationen der Habilitation**