

Erzinger, Andrea B.

Large-Scale Assessments und Forschung zu schulischer Integration

Wicki, Monika T. [Hrsg.]; Törmänen, Minna R. K. [Hrsg.]: *Bildung für alle stärken. Ein Handbuch für die evidenzbasierte Entwicklung inklusiver Schulen.* Bad Heilbrunn : Verlag Julius Klinkhardt 2025, S. 84-93



Quellenangabe/ Reference:

Erzinger, Andrea B.: Large-Scale Assessments und Forschung zu schulischer Integration - In: Wicki, Monika T. [Hrsg.]; Törmänen, Minna R. K. [Hrsg.]: *Bildung für alle stärken. Ein Handbuch für die evidenzbasierte Entwicklung inklusiver Schulen.* Bad Heilbrunn : Verlag Julius Klinkhardt 2025, S. 84-93 - URN: urn:nbn:de:0111-pedocs-332184 - DOI: 10.25656/01:33218; 10.35468/6167-05

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-332184>

<https://doi.org/10.25656/01:33218>

in Kooperation mit / in cooperation with:



<http://www.klinkhardt.de>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden und es darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-Licence: <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work or its contents. You are not allowed to alter, transform, or change this work in any other way.

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



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Andrea B. Erzinger

Large-Scale Assessments und Forschung zu schulischer Integration

Zusammenfassung

Als Large-Scale Assessments (LSA) angelegte Bildungsvergleichsstudien erlauben Aussagen auf der Ebene von Bildungssystemen. Dabei wird die Datengrundlage für Überlegungen zur deren Optimierung (Monitoring) sowie für die Einordnung der Ergebnisse in einen grösseren Kontext (Benchmarking) generiert (Drechsel et al., 2015; Prenzel & Sälzer, 2019; Prenzel et al., 2013). Die Einhaltung diverser Standards schafft die Voraussetzungen dafür, dass LSA valide und reliable Daten liefern (Wagemaker, 2020). In der Folge generieren LSA einen Nutzen für Akteure aus der Bildungspolitik, aber auch die Chance auf detaillierte Beschreibungen und Analysen von Teilpopulationen, wie sie aus der Perspektive verschiedener Forschungsrichtungen gewünscht werden. Gerade im Rahmen der Forschung zu schulischer Integration stellen sich spezifische Herausforderungen für die Nutzung von LSA-Daten. Diese werden in diesem Beitrag aufgezeigt und eingeordnet, wobei auch Möglichkeiten diskutiert werden, wie diesen begegnet werden kann.

Abstract

Large-scale assessments (LSA) in educational comparison studies enable statements at the level of education systems. They provide a data basis for considerations regarding their optimization (monitoring) and for contextualizing the results in a broader context (benchmarking) (Drechsel et al., 2015; Prenzel & Sälzer, 2019; Prenzel et al., 2013). Adherence to various standards ensures that LSAs deliver valid and reliable data (Wagemaker, 2020). Consequently, LSAs generate benefits for education policy actors and offer opportunities for detailed descriptions and analyses of subpopulations from the perspective of various research fields. In the context of research on inclusion, specific challenges arise for the use of LSA data. This article highlights and contextualizes these challenges while also presenting ways to address them.

Advance Organizer

Objective and Overview:

The article “Large-Scale Assessments and Research on School Integration” by Andrea B. Erzinger highlights the role of large-scale assessments (LSA) in educational research and their significance for studying school integration. LSA provide a valuable data foundation for analyzing and optimizing education systems, enabling detailed examination of specific subpopulations, such as students with special educational needs, provided that these were taken into account when sampling.

Key Themes and Concepts:

1. Function and Importance of Large-Scale Assessments (LSA):
 - a) LSA serve to evaluate and compare education systems on a broad scale.
 - b) They provide valid and reliable data that can be used for both educational policy and scientific analysis.
2. Challenges of Using LSA in Research on School Integration:
 - a) Question Formulation: Research questions must be designed so they can be addressed using LSA data.
 - b) Population and Sample Description: The sample must adequately represent the entire population to enable valid conclusions.
 - c) Instrument Development and Delivery: The assessment instruments must be designed to capture data on students with special educational needs.
 - d) Survey Management: Standardized processes and consideration of students with special educational needs are essential for producing comparable results.
3. Benefits and Limitations of LSA Data:
 - a) LSA enable detailed analyses of student performance and its relationships with individual characteristics.
 - b) However, limitations exist for subpopulations of students whose proportion of the analysed population is only partially sufficient for differentiated data analyses, which may apply to students with special educational needs.
4. Strategies to Overcome Challenges:
 - a) Oversampling: This can help adequately capture smaller subpopulations.
 - b) Adaptation and appropriate delivery of tests and questionnaires with regard to specific questions relating to the subpopulation under investigation.
 - c) Linking with External Data Sources: Utilizing additional data sources can help capture relevant indicators.

Connection to Prior Knowledge and Context

Research on school integration provided that suitable questions are formulated and any subpopulations required to answer them are identified. greatly benefits from the insights and data provided by large-scale assessments. Existing knowledge about the challenges and success factors of inclusive education can be expanded through the systematic and comprehensive data collection of LSA provided that suitable accompanying projects are developed. LSA data offer the opportunity to evaluate and optimize the effectiveness of educational measures and systems in an international context. By integrating LSA findings into practice, education systems can become more inclusive and effective, ultimately leading to better learning conditions for all students.

Large-Scale Assessments und Forschung zu schulischer Integration

LSA werfen eine ergebnisorientierte Perspektive auf das Bildungssystem und messen dessen Leistungsfähigkeit anhand der Kompetenzen der Schülerinnen und Schüler (SuS) zu einem bestimmten Zeitpunkt in deren Schulbiografie. Neben der Kompetenzmessung anhand standardisierter Tests erlauben weitere im Rahmen der LSA erfasste Indikatoren u. a. Zusammenhänge von Kompetenzen mit individuellen Merkmalen zu prüfen (z. B. Geschlecht).

Die Berücksichtigung dieser Indikatoren im Untersuchungsdesign und deren adäquate Messung sind zentrale Herausforderungen, damit die Ergebnisse vergleichbar sind und eingeordnet werden können. Darüber hinaus kommt weiteren Standards eine zentrale Bedeutung zu, damit LSA ihren Zweck erfüllen, eine breite und die Population adäquat abbildende Datenbasis für valide und reliable Vergleiche unterschiedlicher Bildungssysteme zu generieren (Wagemaker, 2020). Unter der Berücksichtigung der entsprechenden inhaltlichen, methodologischen und methodischen Anforderungen, erlauben LSA Aussagen über kleinere Teilpopulationen der untersuchten Population und ermöglichen, bestimmte Fragestellungen einer Beantwortung zuzuführen.

Aus der Perspektive der Forschung zu schulischer Integration lassen sich verschiedene Fragestellungen als Anknüpfungspunkte für Analysen von LSA-Daten formulieren. Welche Herausforderungen sich dabei stellen und welche Möglichkeiten zu deren Überwindung sich bieten, wird in diesem Beitrag aufgezeigt. Dabei wird wo angebracht, beispielhaft auf die internationale Vergleichsstudie Programme for International Student Assessment (PISA) referenziert.

Herausforderungen bei der Nutzung von Daten aus LSA für die Forschung zu schulischer Integration

In erster Linie ist sicherzustellen, dass eine Fragestellung formuliert wird, zu deren Beantwortung die Nutzung von Daten aus LSA der adäquate methodische Ansatz darstellt. Dabei gilt es zu berücksichtigen, dass LSA aufgrund ihrer Anlage und Voraussetzungen darauf zielen, die beschriebenen Populationen und bestimmte, aufgrund des Untersuchungsdesigns festgelegte Teilpopulationen zu beschreiben und mit erfassten Kompetenzen und Indikatoren in Zusammenhang zu stellen (Prenzel & Sälzer, 2019).

Die Fragestellung bestimmt,

- (1) welche theoretisch und empirisch relevanten Grundlagen ins Untersuchungsdesign einfließen und bei der Beschreibung der Population und der Stichprobe, der Instrumententwicklung und der Auslieferung sowie des Erhebungsmanagements (Organisation und Durchführung der Erhebungen in den Schulen) berücksichtigt werden müssen,
- (2) welche zentralen Indikatoren und Konstrukte definiert werden müssen, damit die entsprechenden Daten generiert werden und die Frage einer Beantwortung zugeführt werden kann. Weiter ist zentral, die relevanten Indikatoren und Konstrukte vergleichbar zu erfassen, damit deren Einordnung möglich ist.

Beschreibung der Population und der Stichprobe

Kompetenzstudien, die als LSA angelegt sind, beschreiben Populationen von SuS in Bezug auf Kompetenzen (vgl. beispielhaft für PISA: OECD, 2023a). Getestet wird dabei eine die Population adäquat abbildende Stichprobe, damit auf bestimmte Populationsparameter geschlossen werden kann. Um den eingangs erwähnten Zweck der Vergleichbarkeit der Ergebnisse zwischen den Bildungssystemen zu ermöglichen, ist die Standardisierung des Stichprobenziehungsverfahrens von zentraler Bedeutung. Damit wird die Grundlage gelegt, dass entsprechende Schlussfolgerungen valide sind (Prenzel & Sälzer, 2019).

Demnach gilt es bei LSA im Rahmen des Untersuchungsdesigns die Population zu beschreiben, die untersucht werden soll. Bei PISA wird in diesem Zusammenhang geklärt, welche Bildungsinstitutionen Berücksichtigung finden sowie ob und unter welchen Voraussetzungen Ausschlüsse durch einzelne Bildungssysteme erlaubt sind (OECD, 2023b). In der Folge wird nach einem bestimmten Stichprobenziehungsverfahren eine adäquate Stichprobe von zu testenden Bildungsinstitutionen und innerhalb dieser von zu testenden SuS gezogen. Dabei werden Stratifizierungsmerkmale auf der Ebene der Bildungsinstitutionen und auf der Ebene der SuS bestimmt, damit Rückschlüsse von der Stichprobe auf die Population gezogen werden können (siehe auch Verner et al., 2019). Unterschiede lassen sich dann

benennen, wenn von der Stichprobe und den Gruppen auf die jeweilige Zielpopulation geschlossen werden kann. Für weitere Teilpopulationen lassen sich anhand der gezogenen Stichprobe keine gültigen Aussagen treffen (Sälzer, 2015).

Dabei ergibt sich gerade aus dem Anspruch von LSA, dass die Stichprobe die Population adäquat abgebildet soll, die Herausforderung, dass sich einzelne Teilpopulationen von SuS aufgrund der Tatsache, dass ihr Anteil an der untersuchten Population oftmals für differenzierte Datenauswertungen nur bedingt ausreicht, mit LSA-Daten nicht untersuchen lassen (Blanck et al., 2013). So zeigt ein Blick auf den aktuellen Forschungsstand, dass SuS mit besonderem Bildungsbedarf in LSAs international bisher unterrepräsentiert sind (LeRoy et al., 2019; Sälzer et al., 2020), was die Nutzung von LSA-Daten für die Forschung zu schulischer Integration aufgrund der Stichproben, die für LSA gezogen werden, nur beschränkt ermöglicht. Deshalb ist unabdingbar, dass je nach Fragestellung bereits früh im Untersuchungsdesign die Realisierung allfälliger Zusatzstichproben (Oversamplings) mitgedacht wird.

Instrumententwicklung und Auslieferung

Weiter liegen Bildungsvergleichsstudien, die als LSA angelegt sind, einerseits Instrumente zugrunde, welche Kompetenzen zu messen in der Lage sind, andererseits Instrumente, welche Indikatoren erfassen, um den Nutzen der LSA-Daten für Akteure aus Politik und Wissenschaft zu erhöhen. Die Art wie auch der Inhalt der Erhebungsinstrumente, die für LSA entwickelt werden, werden im Rahmen des dem jeweiligen LSA zugrundeliegenden Untersuchungsdesigns auf der Grundlage der theoretischen Rahmenkonzeption definiert (OECD, 2023a).

Damit die aus der Perspektive der Forschung zu schulischer Integration zentralen Fragestellungen mit den Daten beantwortet werden können, sind demnach auf der Ebene des Untersuchungsdesigns die entlang der entsprechenden Fragestellungen relevanten Indikatoren und Konstrukte inhaltlich adäquat zu operationalisieren und angepasst an das standardisierte Erhebungssetting von LSA einzubringen. Dies bezieht sich einerseits auf die Testinstrumente zur Erfassung der Kompetenzen, aber auch auf die Fragebogeninstrumente zur Erfassung weiterer zentraler Indikatoren und Konstrukte. Damit für die Analyse valide und reliable Test- und Fragebogendaten von allen Schülerinnen und Schülern vorliegen, ist zentral, dass bei der Instrumententwicklung auch Tests und Fragebogen für SuS mit besonderem Bildungsbedarf erstellt werden (bei PISA existieren beispielsweise in diesem Zusammenhang Testkurzformen, zudem werden seit PISA 2018 auch adaptive Testdesigns eingesetzt (OECD, 2023b)). Werden unterschiedliche Testformen eingesetzt, gilt es mitzudenken, deren Einsatz und Nutzen im Hinblick auf die Auswertungen und die Vergleichbarkeit der Ergebnisse zu prüfen (Sälzer et al., 2020). Neben der kognitiven ist weiter auch die praktische Komponente zu berücksichtigen. Dabei empfiehlt sich, dass alle LSA-Instrumente barrierefrei

zugänglich sind und die Testauslieferung sowie auch das Erhebungssetting an SuS mit besonderem Bildungsbedarf angepasst sind, damit alle SuS selbstständig und ohne fremde Hilfe die Erhebung absolvieren können (siehe beispielhaft entsprechende Begleitprojekte bei PISA: <https://www.oecd.org/en/about/projects/pisa-rdi-programme.html>).

Erhebungsmanagement

Eine Voraussetzung für valide und reliable Vergleiche von Bildungssystemen stellen bei LSA auch standardisierte Prozesse im Rahmen des Erhebungsmanagements (Organisation und Durchführung der Erhebungen) dar. Bei PISA organisieren in jedem teilnehmenden Land nationale Teams die Erhebungen entlang bestimmter internationaler Standards. Während den Vorbereitungen der Erhebungen stehen die nationalen Teams in engem Kontakt mit den (für die zur Teilnahme an PISA gezogenen SuS) zuständigen Lehrpersonen. Die Durchführung der Erhebungen (Leistungstests sowie Fragebogenuntersuchung) in den Schulen erfolgt jeweils an einem bestimmten Halbtage durch eine dafür geschulte erhebungsadministrierende Person entlang eines standardisierten Erhebungsskripts. Voraussetzung zur Teilnahme der SuS an der Erhebung ist, dass diese in der Lage sind, über die gesamte Dauer der Erhebung selbstständig und ohne individuelle Hilfsmittel teilzunehmen. Im Kontakt mit den Lehrpersonen wird deshalb im Vorfeld der Erhebung geklärt, welche SuS einen allfälligen besonderen Bildungsbedarf haben und ob sie ohne fremde Hilfe an der Erhebung teilnehmen können, demnach als erhebungsfähig eingeschätzt werden (OECD, 2023b), oder von der Erhebung ausgeschlossen werden müssen.

Um LSA-Daten im Rahmen der Forschung zu schulischer Integration zu nutzen, stellt dieser Aspekt je nach Fragestellung, welche mit den Daten angegangen werden soll, eine der grössten Herausforderungen dar. Einerseits ist für die Beschreibung entsprechender Teilpopulationen zentral, dass der besondere Bildungsbedarf einzelner SuS (unabhängig von Teilnahme oder nicht) systematisch und vergleichbar erfasst wird, was ein entsprechendes Kategoriensystem und ein gemeinsames Grundverständnis in Bezug auf dessen Anwendung voraussetzt. Andererseits riskiert die Bestimmung der Teilnahmefähigkeit der SuS durch die jeweilige Lehrperson, dass Ausschlüsse nicht unter vergleichbaren Bedingungen vorgenommen werden und die entsprechenden Teilpopulationen nicht adäquat zwischen den Bildungssystemen verglichen werden können. Gerade bei PISA zeigte sich in der bisherigen Forschung bei den Ausschlussraten eine hohe Variabilität zwischen den Bildungssystemen (LeRoy et al., 2019), was die Vergleichbarkeit der Ergebnisse tangiert. So wurde für Deutschland festgestellt, dass SuS mit besonderem Förderbedarf bisher per se oder individuell eher als nicht erhebungsfähig deklariert und daher von der Teilnahme ausgeschlossen wurden (Müller et al., 2017). Für die Schweiz wurde dieser Aspekt noch nicht vertieft untersucht. Zudem steht

der Vergleichbarkeit dieser Teilpopulationen auch im Weg, dass im Rahmen des Stichprobenziehungsverfahrens aus praktischen Gründen darauf verzichtet wird, ausgeschlossene SuS zu ersetzen. In der Folge riskiert gerade der gehäufte Ausschluss von SuS mit besonderem Bildungsbedarf, dass sogar bei Stichproben mit Oversamplings die Population allenfalls nicht mehr angemessen abgebildet werden kann.

Aus der Perspektive der Forschung zu schulischer Integration zentrale Indikatoren werden teilweise während des Erhebungsmanagements erfasst und riskieren, durch subjektive Einschätzungen verzerrt zu werden. Dem lässt sich mit einem vergleichbaren Kategoriensystem zur Erfassung der zentralen Aspekte im Zusammenhang mit dem besonderen Bildungsbedarf von SuS entgegenwirken (Sahli Lozano et al., 2021). Dabei ist jedoch auch zentral, dass in Bezug auf dessen Anwendung im Zusammenhang mit der Beurteilung der Teilnahmefähigkeit ein gemeinsames Grundverständnis besteht. Hier sind aus der Perspektive der Forschung zu schulischer Integration die Erhöhung der Standardisierung von Ausschlüssen auf der Ebene der SuS sowie der Erhalt von ausreichend grossen Teilpopulationen zur Beantwortung der entsprechenden Fragestellung die zentralen Ziele.

Alternativ zur Erfassung der zentralen Indikatoren im Rahmen der Instrumententwicklung oder während des Erhebungsmanagements besteht auch die Möglichkeit, relevante Indikatoren aus anderen Quellen (z. B. über die Verknüpfung mit administrativen Daten oder Daten aus anderen Registern) zu generieren und diese mit den vorhandenen LSA-Daten zu verknüpfen (siehe beispielhaft Balestra et al., 2022). Je nach Art der Quelle weisen die Indikatoren im Hinblick auf die Beantwortung der Fragestellung in ihrer Natur liegende Vor- und Nachteile auf (Ackermann & Erzinger, 2023). Demnach ist zentral, bereits im Rahmen des Untersuchungsdesigns die zur Beantwortung der entsprechenden Fragestellung relevanten Indikatoren und die dafür validen Quellen zu benennen sowie deren Erschliessung und Verknüpfung mit den Erhebungsdaten frühzeitig mitzudenken. Das ist auch im Hinblick auf die Reduktion der Befragungslast aller Teilnehmenden zentral (once-only Prinzip).

Umgang mit den Herausforderungen im Rahmen von Begleitprojekten

Um Daten aus LSA für die Forschung zu schulischer Integration optimal nutzen zu können, gilt es in erster Linie die aus der spezifischen Forschungsperspektive zentrale Fragestellung zu formulieren und zu prüfen, ob und wie diese vor dem Hintergrund der spezifischen theoretischen, methodologischen und praktischen Herausforderungen des entsprechenden LSA-Untersuchungsdesigns angegangen

werden kann. In einem zweiten Schritt ist zentral, ein entsprechendes Begleitprojekt zu konzeptualisieren und dessen Umsetzung zu planen.

Je nach Voraussetzungen sind dabei folgende Aspekte mitzudenken:

Lässt sich die Fragestellung ohne weitergehende Anpassungen des bestehenden Untersuchungsdesigns beantworten, besteht das Begleitprojekt darin, relevante Indikatoren und Konstrukte inhaltlich adäquat zu operationalisieren, angepasst an das standardisierte Erhebungssetting von LSA einzubringen und in die Daten integrieren zu lassen. Dabei lassen sich Indikatoren über eingesetzte Instrumente wie z. B. Fragebogen erfassen. Darüber hinaus lassen sie sich während des Erhebungsmanagements generieren oder aus externen Quellen mit den LSA-Daten verknüpfen (z. B. über entsprechende Register oder Erhebungsdaten, in der Schweiz u. a. über Datenbestände des BFS möglich).

Zentral im Hinblick auf die Operationalisierung ist, dass bei der zielgruppengerechten und korrekten Erfassung der im Rahmen der Fragestellung zentralen Indikatoren ein vergleichbares Kategoriensystem angewendet wird und auch dessen Anwendung systematisch sowie empirisch begründet erfolgt.

Wird bei der Konzeptualisierung des Begleitprojekts jedoch deutlich, dass sich aus der Perspektive der Forschung zu schulischer Integration in Bezug auf das LSA-Untersuchungsdesign Herausforderungen inhaltlicher, methodologischer oder praktischer Art stellen, gilt es diese frühzeitig anzugehen.

- Herausforderungen im Zusammenhang mit einer Stichprobe, welche die Population zwar adäquat abbildet, aber die für die Beantwortung der Fragestellung relevante Teilpopulationen von SuS nur bedingt abdeckt, erfordern als Oversampling geplante Begleitprojekte. Anhand von Oversamples können bestimmte Teilpopulationen (wie z. B. SuS mit integrierten Massnahmen) differenzierter beschrieben werden, sodass die Ausschlussraten empirisch begründet und für den Vergleich der Bildungssysteme stärker als bisher vereinheitlicht werden können. Dadurch wird die LSA-Datenbasis breiter und belastbarer (dabei wird jedoch das Problem des systematischen Ausschlusses von SuS mit bestimmtem besonderem Bildungsbedarf noch nicht eliminiert). Ein entsprechendes Begleitprojekt würde auch einen bildungspolitischen Mehrwert aufweisen, da die empirische Evidenz als Bezugspunkt bildungspolitischer Entscheidungen belastbarer würde, wenn alle SuS repräsentativ erfasst werden (Sälzer et al., 2020).
- Herausforderungen im Zusammenhang mit dem Stichprobenziehungsverfahren (z. B. aufgrund von auf der Schul- oder der SuS-Ebene angewendeten Stratifizierungsmerkmalen oder von nicht empirisch nachvollziehbaren Ausschlüssen), die das Angehen der Forschungsfrage erschweren, erfordern als erster Schritt im Rahmen der Begleitprojekte gezielte, empirisch begründete Optimierungsvorschläge bei der Umsetzung der Stichprobenziehung.

- Bestehen Herausforderungen auf der Ebene der untersuchten Inhalte, gilt es diese auf der Grundlage der theoretischen Rahmenkonzeption anzugehen, damit zur Beantwortung der Fragestellungen alle nötigen Informationen auf der Ebene der Schule und SuS vorliegen. Gerade in Bezug auf die Instrumente muss weiter darauf geachtet werden, dass diese auch von SuS mit besonderem Bildungsbedarf bearbeitet werden können (bei PISA bietet die OECD beispielsweise eine Testkurzform, sog. *Une Heure (UH) Forms* an (OECD, 2023b), an welcher die Schweiz sich bisher jedoch jeweils nicht beteiligt hat). Über die kognitive Komponente hinaus, gilt es hier auch die praktische zu berücksichtigen. Die ausgelieferten Instrumente sollten möglichst barrierefrei zugänglich sein und die Organisation der Erhebungen den SuS mit besonderem Bildungsbedarf oder anderen für die Beantwortung der entsprechenden Fragestellung relevanten SuS angepasst (vgl. beispielhaft für PISA: OECD, 2024).

Abschliessend gilt es mit Blick auf bildungspolitische Herausforderungen zu empfehlen, entsprechende Begleitprojekte in einem ersten Schritt im Kleinen zu implementieren (z. B. über wenige Länder oder in föderalistischen Bildungssystemen über wenige Regionen hinweg), bevor deren flächendeckende Einführung geplant wird, so dass entsprechende Ressourcen gebündelt und erste Teilerfolge genutzt werden können, um breitere Kreise zu erreichen.

Literatur

- Ackermann, T. & Erzinger, A. B. (2023). *Evaluation von Datenverknüpfungen mit Daten aus dem Scientific Use File der ÜGK und Registerdaten des BFS*. Universität Bern: Interfaculty Centre for Educational Research.
- Balestra, S., Eugster, B., & Liebert, H. (2022). Peers with Special Needs: Effects and Policies. *The Review of Economics and Statistics*, 104(3), 602–618. https://doi.org/10.1162/rest_a_00960
- Blanck, J. M., 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. <https://doi.org/10.5169/seals-814996>
- Drechsel, B., Prenzel, M., & Seidel, T. (2015). Nationale und internationale Schulleistungsstudien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (343–368). Springer Berlin Heidelberg.
- Gresch, C., Kuhl, P., Grosche, M., Sälzer, C., & Stanat, P. (Hrsg.). (2020). *Schüler*innen mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Einblicke und Entwicklungen*. Springer VS. <https://doi.org/10.1007/978-3-658-27608-9>
- LeRoy, B. W., Samuel, P., Deluca, M., & Evans, P. (2019). Students with special educational needs within PISA. *Assessment in Education: Principles, Policy & Practice*, 26(4), 386–396. <https://doi.org/10.1080/0969594X.2017.1421523>
- Müller, K., Prenzel, M., Sälzer, C., Mang, J., Heine, J.H., & Gebhardt, M. (2017). Wie schneiden Schülerinnen und Schüler an Förderschulen bei PISA ab? Analysen aus der PISA 2012-Zusatzerhebung zu Jugendlichen mit sonderpädagogischem Förderbedarf. *Unterrichtswissenschaft*, 45(2), 175–192. <https://doi.org/10.3262/UW1702175>
- OECD. (2023a). *PISA 2022 Assessment and Analytical Framework*. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>

- OECD. (2023b). *PISA 2022 Technical Report*. <https://www.oecd.org/pisa/data/pisa2022technicalreport/>
- OECD. (2024). *Mapping Study for the Integration of Accommodations for Students with Special Education Needs (SEN) in PISA*. Education Working Paper. [https://one.oecd.org/document/EDU/WKP\(2024\)03/en/pdf](https://one.oecd.org/document/EDU/WKP(2024)03/en/pdf)
- Prenzel, M. & Sälzer, C. (2019). Large-scale assessments of educational systems. In R. Becker (Hrsg.), *Research handbook on the sociology of education* (S. 536–552). Edward Elgar publishing.
- Prenzel, M., Sälzer, C., Klieme, E., & Köller, O. (Hrsg.). (2013). *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Waxmann.
- Sahli Lozano, C., Cramer, S., & Adeifo Gosteli, D. (2021). *Integrative und separative schulische Massnahmen in der Schweiz (InSeMa): Kantonale Vergabe- und Umsetzungsrichtlinien*. www.szh-csp.ch/b2021-01
- Sälzer, C. (2015). *Erkenntnisse aus Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen* [Habilitation]. TUM School of Education, München.
- Sälzer, C., Denk, A., Müller, F., & Grosche, M. (2020). Spezifische Herausforderungen des Einbezugs von Schüler*innen mit sonderpädagogischem Förderbedarf in internationale Large-Scale Assessments. In C. Gresch, P. Kuhl, M. Grosche, C. Sälzer & P. Stanat (Hrsg.), *Schüler*innen mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Einblicke und Entwicklungen* (S. 55–74). Springer VS.
- Verner, M., Erzinger, A. B., & Fässler, U. (2019). Zur Schweizer Stichprobe PISA 2015. Eine externe Validierung zentraler Stichprobenmerkmale. *Swiss Journal of Educational Research*, 41(2), 522–542. <https://doi.org/10.24452/sjer.41.2.15>
- Wagemaker, H. (2020). Study design and evolution, and the imperatives of reliability and validity. In H. Wagemaker (Hrsg.), *Reliability and Validity of International Large-Scale Assessment* (S. 7–21). Springer International Publishing.

Author

Erzinger, Andrea B. is Managing Director of the Interfaculty Centre of Educational Research (ICER) at the University of Bern and National Project Manager (NPM) of PISA (Programme for International Student Assessment) Switzerland