

Open Science als Weg zu verlässlicher sonderpädagogischer Forschung

Timo Lüke^{1,2}

¹ Karl-Franzens-Universität Graz

² Forschungszentrum für Inklusive Bildung (FZIB)

Zusammenfassung: Im Rahmen der Glaubwürdigkeits- oder Replikationskrise sehen sich Forschende aktuell mit der Forderung nach mehr Transparenz in ihrer Arbeit und teilweise sogar dem Vorwurf mangelnder Sorgfalt konfrontiert. Prinzipien, die Forschende als Teil ihres Berufsethos verstehen dürften, scheinen in der alltäglichen Arbeit und in institutionalisierten Prozessen nicht immer den höchsten Stellenwert zu haben. Open-Science-Praktiken sind Arbeitsweisen, die Forschende dabei unterstützen, den gesamten Forschungsprozess transparenter sowie die Kommunikation mit der Forschungsgemeinschaft offener und unverzüglicher zu gestalten. Die Art und Weise, wie sich wissenschaftliche Kommunikation in den vergangenen Jahrzehnten verändert hat und vermutlich weiterhin verändern wird, macht eine offene, transparente und unverzügliche Kommunikation mit der Forschungsgemeinschaft immer einfacher. Der Zeitpunkt, diese neuen Möglichkeiten im Sinne besserer wissenschaftlicher Praxis zu nutzen, ist daher günstig.

Schlüsselbegriffe: Open Science, Sonderpädagogik, Replikationskrise, Metawissenschaft, Forschungsmethoden

Open Science as a Path to Reliable Special Education Research

Summary: In the context of the credibility or replication crisis, researchers are currently confronted with demands for more transparency in their work and sometimes even accusations of a lack of due diligence. Principles that researchers are likely to understand as part of their professional ethos do not always seem to have the highest priority in everyday work and institutionalized processes. Open science practices help researchers make the entire research process more transparent and the communication with the research community more open and immediate. The way scientific communication has changed over the past decades and is likely to continue to change makes open, transparent, and immediate communication with the research community increasingly easy. The time is, therefore, right to take advantage of these new opportunities in the interest of better scientific practice.

Keywords: Open science, special education, replication crisis, metascience, research methods

“Open Science is just science done right.”
(J. P. Tennant)

In den vergangenen zehn Jahren haben die überwiegend erfolglosen Versuche, sozialwissenschaftliche Studien, die in hochwertigen Fachzeitschriften (insbesondere in der Psychologie) publiziert wurden, zu replizieren, für Auf-

sehen gesorgt. Ein großer Teil der dort kommunizierten (häufig viel beachteten) Befunde ließ sich unter vergleichbaren Bedingungen nicht wiederholen. In anderen Studien konnten nicht einmal auf der Basis der gleichen Datensätze die ursprünglichen Ergebnisse reproduziert werden. Die Replikations- oder Glaubwürdigkeitskrise schlug und schlägt innerhalb und

außerhalb des Wissenschaftssystems zu Recht große Wellen. Ein viel diskutierter Weg aus dieser Krise sind Open-Science-Praktiken wie Open Access, Open Data oder Präregistrierung. Im vorliegenden Beitrag wird gezeigt, dass Open Science-Praktiken im Wesentlichen das widerspiegeln, was die meisten Forschenden ohnehin als gute wissenschaftliche Praxis bezeichnen würden. Daher könnte die aktuelle Entwicklung in Richtung Open Science vielleicht als wissenschaftspolitische Kurskorrektur bezeichnet werden, die fehlgeleitete Anreizsysteme und institutionalisierte Prozesse wieder mit immer dagewesenen Grundwerten vereinbaren soll.

Es gibt mehrere umfassende Argumentationsstränge, anhand derer Open-Science-Praktiken begründet werden. Viele davon werden hier außen vor gelassen. Nicht weil sie weniger überzeugend wären, sondern schlicht wegen der Kürze und Pointiertheit des Formates. Unter anderem argumentieren Open-Science-Unterstützer/innen, dass 1) wissenschaftliche Arbeiten und die daraus gewonnenen Erkenntnisse letztlich von allen Bürger/innen finanziert werden und dementsprechend auch der Zugang (und die Verwertungsrechte) der Allgemeinheit zustünden oder 2) viele Forschende und Institutionen wegen der unglaublichen Preispolitik einiger Großverlage keinen Zugriff auf relevantes Wissen mehr hätten und so systematisch benachteiligt würden.

Im vorliegenden Essay soll vor allem über die Qualität der alltäglichen Arbeit einer jeden Wissenschaftlerin und eines jeden Wissenschaftlers und dem so geschaffenen Wissen argumentiert werden. Dahinter steht die Annahme, dass alle Forschenden in unserer wissenschaftlichen Community verlässliche Ergebnisse erzeugen möchten, um die Allgemeinheit mit dem bestmöglichen Wissen – im Fall der Heilpädagogik und ihrer Nachbargebiete zu Fragen rund um Menschen mit Behinderungen – zu versorgen. Unsere Priorität ist also die Suche nach Wahrheit. Es soll gezeigt werden, dass

wissenschaftliche Grundwerte, die vermutlich alle Forschenden teilen, in den aktuellen institutionalisierten Praktiken vernachlässigt werden und dass insbesondere durch die breite und niedrigschwellige Verfügbarkeit neuer digitaler Kommunikations- und Kollaborationswerkzeuge eine Rückkehr zu einer wissenschaftlichen Praxis, die diesen Grundwerten entspricht, leicht umsetzbar ist.

Mertons Normen und das wissenschaftliche Ethos

Als die vier mertonischen Normen *Kommunismus*, *Universalismus*, *Desinteressiertheit* und *Organisierte Skepsis* sind institutionelle Grundsätze bekannt, die Merton in *The Sociology of Science* beschrieb (Merton, 1973). Sie sind sein Versuch, das wissenschaftliche Ethos – auch im Widerspruch zu systemimmanenten Belohnungssystemen – zu beschreiben und in institutionelle Imperative zu fassen. Die Grundsätze werden auch als CUDO-Normen [Communism, Universalism, Desinterestedness, Organized Scepticism] – später in leicht veränderter Form als CUDOS-Normen [Communism, Universalism, Desinterestedness, Originality, Scepticism] – bezeichnet. Im Folgenden soll vor allem auf die Prinzipien *Kommunismus* und *Desinteressiertheit* eingegangen werden; auch die beiden anderen Normen (*Universalismus* und *Organisierte Skepsis*) sind aber mit den hier geforderten Praktiken einer offenen Wissenschaft sehr gut vereinbar.

Mertons Norm *Kommunismus* (ebd., S. 273) beschreibt das gemeinschaftliche Eigentum an allen wissenschaftlichen Gütern. Zwar verdiene die Wissenschaftlerin oder der Wissenschaftler Anerkennung und Wertschätzung für die eigene Arbeit; die eigentlichen Erkenntnisse seien aber immer geistiges Eigentum der gesamten Forschungsgemeinschaft. „Geheimhaltung ist das Gegenteil dieser Norm, vollständige und offene Kommunikation ihre Umsetzung“ (ebd.,

S. 274). So werde eine zielgerichtete und ertragreiche Kooperation befördert und letztlich das Ziel der wissenschaftlichen Profession insgesamt verfolgt. In seiner Darstellung der Norm *Desinteressiertheit* greift Merton diese Grundannahme auf und argumentiert, dass wissenschaftliche Institutionen und Forschende immer zum Vorteil des gemeinsamen wissenschaftlichen Unternehmens und nicht zum persönlichen Vorteil einzelner Personen arbeiteten. In keiner anderen Profession sei das Bedürfnis nach Überprüfbarkeit durch fachkundige Peers so groß wie in der Wissenschaft und nirgends sei auch die tatsächliche Kontrolle so groß.

Zusammenfassend lässt sich festhalten, dass zum institutionellen Imperativ der Wissenschaft im Sinne Mertons nicht nur die vollständige und offene Kommunikation über alle Arbeitsschritte und gewonnenen Erkenntnisse gehört. Er spricht den Einzelnen sogar jede Berechtigung für Besitzansprüche an Methoden und Erkenntnissen ab.

Replikationsstudien sind unverzichtbar

Im Folgenden betrachten wir die Ebene der tatsächlichen wissenschaftlichen Arbeit und der Entstehung neuen Wissens bzw. der empirischen Bewährung einer wissenschaftlichen Annahme. Ein Ergebnis muss wiederholbar sein, um als wissenschaftliche Erkenntnis zu gelten. „Sogar unsere eigenen Beobachtungen pflegen wir wissenschaftlich nicht ernst zu nehmen, bevor wir sie nicht selbst durch wiederholte Beobachtungen oder Versuche nachgeprüft und uns davon überzeugt haben, daß [sic!] es sich nicht nur um ein einmaliges, zufälliges Zusammentreffen handelt, sondern um Zusammenhänge, die durch ihr gesetzmäßiges Eintreffen, durch ihre Reproduzierbarkeit grundsätzlich intersubjektiv nachprüfbar sind“ (Popper, 1935, S. 17). Im Idealfall werden die Beobachtungen des einen Wissenschaftlers also von einer anderen Wissen-

schaftlerin unabhängig repliziert. Das Vorliegen wenigstens einer zweiten Studie – vorzugsweise aus einer anderen Arbeitsgruppe – wird auch in der sonderpädagogischen Forschung als Kriterium zur Beurteilung der Wirksamkeit von Fördermaßnahmen oder Unterrichtsmethoden herangezogen (Wong et al., 2015). Auch Fingerle und Ellinger forderten von *bewährten* (Kategorie 1) und *vermutlich effektiven* (Kategorie 2) Förderansätzen das Vorliegen mehrerer empirischer Studien und verwiesen Methoden, die bisher nur in *einer einzigen* Studie untersucht wurden, in die letzte Kategorie: *potenziell effektiv* (Fingerle & Ellinger, 2008, S. 9).

Sehen wir uns eine fiktive Studie an, die weitgehend als Idealfall einer Wirksamkeitsstudie (am Beispiel einer Leseförderung) gesehen werden kann: „Die Kinder der Stichprobe entsprechen der Zielgruppe, die Lernvoraussetzungen (...) sind differenziert diagnostiziert, das Leseprogramm ist so detailliert dokumentiert, dass man es wiederholen (...) könnte, die Aktionen der Lehrpersonen wurden protokolliert, die Lesefortschritte der Kinder sind nachvollziehbar ermittelt worden und es gibt eine Kontrollgruppe mit vergleichbaren Lernenden und unter vergleichbaren Bedingungen ...“ (Wember, 2020, S. 233f.).

Nehmen wir nun weiterhin an, dass die Gruppe, die diese neue Leseförderung erhielt, später bedeutend besser lesen kann, deutet dies darauf hin, dass die Leseförderung (in dieser Zielgruppe) wirksam ist. Wember (ebd., S. 234) hält eine Empfehlung dieses Trainings als *nachweislich wirksame* oder *evidenzbasierte Methode* aber für vorschnell. Erst wenn es gelingt, den förderlichen Effekt unter anderen Bedingungen (z. B. mit anderen Kindern und anderen Lehrkräften) zu replizieren – also durch die wiederholte empirische Bewährung –, dürfen wir Vertrauen in die Wirksamkeit dieser Leseförderung entwickeln und diese dann auch für den Einsatz in den Schulen empfehlen (Fingerle & Ellinger, 2008; Wember, 2020; Wong et al., 2015).

Selbst wenn wir unserer gesamten Forschungsgemeinschaft höchste Redlichkeit unterstellen, obwohl wir davon ausgehen müssen, dass wenigstens fragwürdige Forschungspraktiken (*questionable research practices*) wie *HARKing* und *p-hacking* auch in der empirischen Sonderpädagogik vorkommen (Makel, Hodges, Cook & Plucker, 2021), sollten wir darauf bestehen, dass unsere Erkenntnisse von uns und anderen kritisch hinterfragt werden und sich letztlich auch durch Replikationsversuche empirisch bewähren müssen.

Fehlende Replikationsstudien sind ein Problem

Makel und Kolleg/innen untersuchten mehr als 45.000 Publikationen aus 36 sonderpädagogischen Fachzeitschriften und erfassten den Anteil der Replikationsstudien. Lediglich 0,5 % der bis dahin veröffentlichten Studien in den sonderpädagogischen Fachzeitschriften sind demnach Replikationsstudien (Makel et al., 2016, S. 208). Damit scheint der Anteil von Replikationsstudien mit dem in benachbarten Disziplinen wie den Erziehungswissenschaften mit circa 0,1 % (Makel & Plucker, 2014) und der Psychologie mit circa 1 % (Makel, Plucker & Hegarty, 2012) vergleichbar zu sein. Was bedeutet das aber für unser Vertrauen in die Arbeiten, die wir als den aktuellen Forschungsstand bezeichnen? Ein verschwindend geringer Teil dieser Befunde wurde nachweislich unabhängig wiederholt.

Ein kleiner Hoffnungsschimmer ist die vergleichsweise hohe „Erfolgsquote“ (Anteil der Replikationsstudien, die aus Sicht der Autor/innen *erfolgreich* die Ergebnisse der Originalstudie repliziert haben) dieser sehr seltenen Replikationsstudien in der Sonderpädagogik. Sie liegt mit 80 % im Vergleich zu anderen Disziplinen recht hoch. Etwas getrübt wird

diese Hoffnung dadurch, dass (1) die Anzahl der Replikationsstudien verschwindend gering ist, (2) nicht erfolgreiche Replikationsstudien wahrscheinlich deutlich seltener publiziert werden und (3) bei etwa zwei Dritteln der Replikationen wenigstens ein/e Autor/in der Originalstudie beteiligt war, was mit höheren „Erfolgsquoten“ assoziiert ist (Makel et al., 2016).

Bedeutet das nicht, dass wir – die oben formulierten Anforderungen an evidenzbasierte Unterrichtsmethoden im Hinterkopf – zugeben müssen, dass die empirische sonderpädagogische Forschung bisher nur wenig hervorbringt, das diesem hohen Anspruch genügt? In aller Regel scheinen wir einen einmaligen Befund als ausreichenden „Beweis“ für ein Studienergebnis – beispielsweise die Wirksamkeit unserer neuen Leseförderung – zu sehen. Und dabei gehen wir hier ja noch davon aus, dass jede einzelne dieser Studien quasi fehlerfrei geplant und durchgeführt wurde und ganz eindeutige Ergebnisse hervorbrachte. Das dürfte in der Realität selten der Fall sein.

Gründe für den Mangel an Replikationsstudien

Ein wichtiger Grund für diesen fächerübergreifenden Mangel an Replikationsstudien ist recht offensichtlich: Einer Replikationsstudie wird in der Regel eine geringere Bedeutung oder Relevanz zugesprochen als einer neuen Studie mit vergleichbarem Aufwand. Dies steht im deutlichen Widerspruch zur oben formulierten Annahme, dass wir wissenschaftlich nur das ernst nehmen, was sich wiederholt und unabhängig beobachten lässt.

Ein weiteres Hindernis einer gesunden Replikationskultur sind die damit verbundenen Herausforderungen und Risiken im Forschungs-

prozess: Damit ein Ergebnis wiederholbar oder reproduzierbar ist, muss der gesamte Forschungsprozess der Originalstudie so transparent gemacht worden sein, dass andere die Studie wiederholen können. Dafür sind sehr detaillierte Informationen unter anderem über das methodische Vorgehen oder die Auswertung und Interpretation der Resultate notwendig. Zwar wird in vielen Publikationen so viel Transparenz wie möglich hergestellt. Aber alle Wissenschaftlerinnen und Wissenschaftler, die jemals versucht haben, eine Studie zu replizieren, werden zustimmen, dass über die Angaben in einer etwaigen Publikation hinaus noch viele weitere Informationen notwendig waren, um die Bedingungen der Originalstudie annähernd wiederherstellen zu können.

Unzählige Male habe ich in Publikationen gelesen, dass die Materialien und Daten einer Studie bei Bedarf von der Erstautorin oder dem Erstautor angefordert werden könnten. Es ist nicht nur eine persönliche Erfahrung, dass solche Anfragen – werden sie denn einmal gestellt – nur selten *überhaupt* beantwortet werden und auch dann so sparsam wie möglich. Nur ein Teil der Forschenden teilt die eigenen Materialien und Daten wirklich mit der Forschungsgemeinschaft (Lüke, Raich, Barth & Steiner, in Vorb.; Makel et al., 2021). Böse Absichten sollen den betreffenden Kolleginnen und Kollegen nicht unterstellt werden. Aber steigern solche Erfahrungen das Vertrauen in die betreffenden Arbeiten und in die Belastbarkeit der daraus hervorgegangenen Ergebnisse?

Was bedeutet es für die empirische Bewährung der fiktiven Leseförderung, wenn die Autor/innen den statistisch bedeutsamen Gruppenunterschied in ihrer Studie in einem Forschungsbericht veröffentlichen, die Daten, die verwendeten Materialien, die Hinweise für Lehrkräfte usw. aber nicht verfügbar gemacht werden?

Wenn wir gemeinsam auf der Suche nach Wahrheit oder der bestmöglichen Annäherung an Wahrheit sind, sollte es in unserem Interesse sein, alles, was Kolleg/innen zur Replikation unserer Studien benötigen, ohne Hürden frei verfügbar zu machen. Tun wir dies nicht, verabschieden wir uns vom oben beschriebenen Berufsethos und verhindern aktiv, dass unsere Ergebnisse kritisch geprüft und repliziert werden können. Und wie bereits ausgeführt ist dies eine zentrale Voraussetzung dafür, dass die Autor/innen ihre Leseförderung unter bestimmten Bedingungen überhaupt als wirksame Methode oder evidenzbasierte Praxis bezeichnen können.

Open-Science-Praktiken als Weg aus der Replikations- und Glaubwürdigkeitskrise

Innerhalb der wissenschaftlichen Gemeinschaft sollten Erkenntnisse und Methoden als Gemeingut angesehen werden. Die Kommunikation über alle Schritte wissenschaftlicher Arbeit sollte vollständig und offen erfolgen und genau dafür sollten Wissenschaftlerinnen und Wissenschaftler Anerkennung erfahren, nicht durch das vermeintliche Eigentum an bestimmten Methoden oder Daten oder für ihre schriftliche Kommunikation. Wissenschaftlerinnen und Wissenschaftler und mit ihnen wissenschaftliche Institutionen haben sich in ihrer täglichen Arbeit und in institutionalisierten Entscheidungen über viele Jahre hinweg von einigen ihrer Grundprinzipien entfernt. Open-Science-Praktiken bestärken Forschende darin, in ihrer täglichen Arbeit wieder mehr Wert auf Offenheit und Transparenz zu legen und institutionelle Prozesse und Entscheidungen im Sinne dieser zu verändern. So sind sie letztlich nichts anderes als die Umsetzung jener Normen, die Merton (und in ähnlicher Form natürlich auch andere) für die Wissenschaft und die Gemeinschaft der For-

schenden gefordert haben: Open Access führt zu einer grundsätzlich besseren Verfügbarkeit von Forschungsergebnissen, Open Data und Open Materials machen die angewendeten Methoden und Ergebnisse der gesamten Forschungsgemeinschaft zugänglich, Präregistrierungen führen zu einem transparenteren und offeneren Umgang mit dem Forschungsprozess. Dieser offene Austausch und diese Transparenz sind letztlich auch Wege aus der Glaubwürdigkeitskrise, die aktuell alle Forschenden beschäftigen sollte, auch wenn die Heilpädagogik und ihre Nachbarggebiete bisher nicht so stark von ihr erschüttert wurden wie manche Nachbardisziplin. Die Glaubwürdigkeit aller Wissenschaften ist gefährdet und muss geschützt werden.

Historisch wurde diese Kommunikation innerhalb einer noch sehr überschaubaren Forschungsgemeinschaft zunächst in Briefform realisiert, später gründete man Akademien und andere Organisationen, die Konferenzen und Fachzeitschriften als Kommunikationsplattformen etablierten. Diese Formate veränderten sich in der Vergangenheit sehr stark. Auch aktuell verändern die zunehmende Digitalisierung und Globalisierung die Kommunikationsformen und Arbeitsweisen von Wissenschaftlerinnen und Wissenschaftlern deutlich. Es ist davon auszugehen, dass viele dieser Kommunikationsformate in absehbarer Zeit abgelöst werden. Die Ausgangslage für die oben beschriebenen Veränderungen könnte also nicht besser sein, als sie es aktuell ist: Die Kollaboration mit Forschenden in aller Welt ist so niedrigschwellig möglich wie noch nie. Kommunikation und gemeinsames Arbeiten können in Echtzeit erfolgen. Auch die Weitergabe neuer Ideen, Überarbeitungen früherer Arbeitsschritte, Verbesserungen von Materialien oder Analysen sind heute viel leichter transparent und uneingeschränkt zu kommunizieren. Lösungen, um Forschenden Anerkennung für ihre transparente Arbeitsweise zu geben, wurden bereits etabliert und werden

von immer mehr Forschungseinrichtungen und Universitäten in Entscheidungsprozesse einbezogen.

Dass immer mehr Forschende Open-Science-Praktiken anwenden, dass immer mehr Wissenschaftsorganisationen und Drittmittelgeber Open Science von ihren Mitarbeitenden und Geförderten einfordern sowie auf wissenschaftspolitischer Ebene (z. B. UNESCO, ERC, DFG, SNE, FWF) durchsetzen, ist keine Neuerfindung wissenschaftlichen Arbeitens. Es ist eine Rückbesinnung auf Grundprinzipien guter wissenschaftlicher Praxis.

Literatur

- Fingerle, M. & Ellinger, S. (2008). Sonderpädagogische Förderprogramme im Vergleich. In M. Fingerle & S. Ellinger, *Sonderpädagogische Förderprogramme im Vergleich: Orientierungshilfen für die Praxis*, 7–10. Stuttgart: Kohlhammer.
- Lüke, T., Raich, K., Barth, M. & Steiner, T. M. (in Vorbereitung). *Open Science-Praktiken in der empirischen sonderpädagogischen Forschung*.
- Makel, M. C., Plucker, J. A. & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7(6), 537–542. <https://doi.org/10.1177/1745691612460688>
- Makel, M. C. & Plucker, J. A. (2014). Facts are more important than novelty. *Educational Researcher*, 43(6), 304–316. <https://doi.org/10.3102/0013189X14545513>
- Makel, M. C., Plucker, J. A., Freeman, J., Lombardi, A., Simonsen, B. & Coyne, M. (2016). Replication of special education research: Necessary but far too rare. *Remedial and Special Education*, 37(4), 205–212. <https://doi.org/10.1177/0741932516646083>
- Makel, M. C., Hodges, J., Cook, B. G. & Plucker, J. A. (2021). Both questionable and open research practices are prevalent in education research. *Educational Researcher*, 50(8), 493–504. <https://doi.org/10.3102/0013189X211001356>
- Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: The University of Chicago Press.
- Popper, K. (1935). *Logik der Forschung*. Wien: Springer.

- Wember, F. B. (2020). Pädagogische Reflexion statt Empirischer Evidenz? Fragen eines lesenden Lehrers. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 89(4), 231–237. <https://doi.org/10.2378/vhn2020.art33d>
- Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S. et al. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders*, 45(7), 1951–1966. <https://doi.org/10.1007/s10803-014-2351-z>

Anschrift des Autors

Prof. Dr. Timo Lüke
Karl-Franzens-Universität Graz
Umwelt-, Regional- und
Bildungswissenschaftliche Fakultät
Institut für Bildungsforschung
und PädagogInnenbildung
Merangasse 70/II
A-8010 Graz
E-Mail: timo.lueke@uni-graz.at