

Editorial 2020

Katrin Auspurg* und Jonas Recker

Mehr Offenheit in der Forschung? Eine Evaluation von Open Science Maßnahmen bei der *Zeitschrift für Soziologie*

<https://doi.org/10.1515/zfsoz-2020-0001>

Die *Zeitschrift für Soziologie* (ZfS) zeichnet sich durch Offenheit für die ganze Bandbreite an Soziologie aus (in inhaltlicher, theoretischer und methodischer Hinsicht), vorausgesetzt die Begutachtung der Beiträge durch *Peers* fällt hinreichend positiv aus. Zugleich ist es Anspruch der Zeitschrift, dass sie regelmäßig Maßnahmen zur weiteren Verbesserung der wissenschaftlichen Qualität diskutiert und umsetzt. Gerade das jährlich erscheinende Editorial ist Anlass, die Publikationspraxis kritisch zu reflektieren.

In dieser Tradition steht auch das diesjährige Editorial, mit dem wir uns zum Ziel gesetzt haben, die *Open Science* Praktiken und speziell die offene Bereitstellung von Materialien zu betrachten, die publizierten Analysen mit quantitativen Daten zugrunde liegen. (Die qualitative Forschung nehmen wir zunächst von der Betrachtung aus, wir kommen aber später noch auf sie zurück.) Im Folgenden werden wir zunächst den Hintergrund schildern, um dann Auswertungen zur Bereitstellung und Nachnutzung von Daten zu Aufsätzen in der *Zeitschrift für Soziologie* zu präsentieren.

Wozu mehr Offenheit? Hintergrund der Standards bei der *Zeitschrift für Soziologie*

Nachbardisziplinen wie die Psychologie, Ökonomie und Politikwissenschaft haben in den letzten Jahren in Reaktion auf die so genannte „Glaubwürdigkeitskrise“ in

der Wissenschaft intensiv über eine bessere Offenlegung von Materialien für die Reproduktion und Replikation von Forschung diskutiert. Ursachen für die mangelnde Reproduzierbarkeit von Ergebnissen sind insbesondere das „Cherry Picking“ von signifikanten, Hypothesen konformen Ergebnissen sowie Fehler bei der Aufbereitung und Auswertung von Daten. Als zentrale Korrekturmöglichkeit gelten Reproduktionen von Analysen durch unabhängige Wissenschaftler*innen (siehe dazu z. B. Christensen & Miguel 2018; Wuttke 2019). Bereits Merton (1973 [1942]) hat angemerkt, dass ein „organisierter Skeptizismus“, also das Prüfen von Ergebnissen durch *Peers*, für den wissenschaftlichen Erkenntnisfortschritt essenziell ist. Diese Prüfung setzt allerdings voraus, dass die zugrunde liegenden Materialien zugänglich sind. Wie sich in den letzten Jahren gezeigt hat, ist dies oft nicht der Fall (für Analysen zur Soziologie: Zenk-Möltgen et al. 2018). Um hier Abhilfe zu schaffen, wurde inzwischen in vielen Disziplinen die Bereitstellung von Daten, Auswertungsfiles und allen sonstigen Materialien verpflichtend gemacht, die für den kompletten Nachvollzug publizierter Ergebnisse erforderlich sind (s. dazu z. B. Christensen & Miguel 2018; Mellor 2018).

Auch in der Soziologie mehren sich die Hinweise, dass eine umfassendere Prüfung der Robustheit von Ergebnissen für den wissenschaftlichen Erkenntnisfortschritt wesentlich ist (Christensen et al. 2019). Dies sind mitunter vielfältige *Errata* und Korrekturen. So mussten gleich mehrere Artikel in dem „Flaggschiff-Journal“ *American Sociological Review* (ASR) wegen falschen Kodierungen von fehlenden Werten korrigiert werden.¹ Diese Fehlkodierungen haben zu teils massiven Verzerrungen der Ergebnisse geführt. Unabhängige Nachprüfungen erscheinen zudem deshalb ratsam, da manche der korrigierten Aufsätze mit starken *Policy*-Empfehlungen versehen waren. Dass

*Korrespondenzautorin: Prof. Dr. Katrin Auspurg, LMU München, Institut für Soziologie, Quantitative Methoden der empirischen Sozialforschung, Konradstraße 6, 80801 München, E-Mail: katrin.auspurg@lmu.de

Dr. Jonas Recker, GESIS – Leibniz-Institut für Sozialwissenschaften, Abt. Datenarchiv für Sozialwissenschaften, Unter Sachsenhausen 6–8, 50667 Köln, E-Mail: jonas.recker@gesis.org

¹ Uns bekannt sind diesbezüglich die Aufsätze von Jasso 1985, Herring 2009 und Munsch 2015. In den letzten Ausgaben von ASR gab es zudem weitere Korrekturen, etwa aufgrund von übersehenen extremen Datenpunkten (für Beispiele siehe auch Christensen et al. 2019; Munoz & Young 2018).

wir hier auf anekdotische Evidenz verweisen müssen, ist dabei einem strukturellen Problem geschuldet: Aktuell scheitert bereits die systematische Prüfung, wie robust soziologische Analysen sind, an einem fehlenden umfassenden Datenzugang für Re-Analysen.

Die möglichst transparente Offenlegung von Forschungsmethoden ist neben Fehlerkorrekturen in einer multi-paradigmatischen Disziplin wie der Soziologie aber auch deshalb geboten, um *Peers* einen klaren Nachvollzug der Ergebnisse und Schlussfolgerungen zu ermöglichen. Gerade wenn unkonventionell vorgegangen wird, sind ausführliche Dokumentationen wichtig, um das methodisch-analytische Vorgehen intersubjektiv verstehen zu können. Entsprechend zählen eine möglichst präzise Beschreibung der Methoden und Forschungsdesigns zu den einschlägigen Richtlinien für „gute soziologische Praxis“ (siehe den Ethik-Kodex der Deutschen Gesellschaft für Soziologie 2017; sowie die Richtlinien zur Datennachnutzung der Akademie für Soziologie 2019).

In quantitativen Auswertungen wird maximale Transparenz in der Regel erst durch die Bereitstellung von Daten und Auswertungsfiles erzielt. Aufgrund der nahezu unzähligen Möglichkeiten, in großen Datensätzen mit nicht-experimentellen Daten Fälle zu selektieren, Variablen zu kodieren, oder auch Auswertungsmodelle zu spezifizieren (siehe zu diesem „garden of forking paths“ Gelman & Loken 2003; Munoz & Young 2018), bieten üblicherweise nur die Auswertungsfiles die Chance, Ergebnisse vollständig nachvollziehbar zu dokumentieren. Idealerweise besteht zugleich Datenzugang, mit dem Nutzer*innen sich selbst von der Robustheit der Ergebnisse im Hinblick auf selbst spezifizierte Variationen der Auswertungsmethoden überzeugen können. Das Angebot an mögliche Kritiker*innen, Ergebnisse selbst zu prüfen, ist wohl das beste Argument, um mögliche Zweifel an der Tragfähigkeit wissenschaftlicher Erkenntnisse auszuräumen (Merton 1973 [1942]).² Die Möglichkeit des Nachvollzugs erhöht zudem die Chancen für den Erkenntnisfortschritt durch kumulative Forschung (Ioannidis 2005): Erst die umfassende Kenntnis von Methoden erlaubt es, bestehende Forschung durch *gezielte* Variationen im Forschungsdesign zu erweitern. Bei (international) vergleichender Forschung lassen sich etwa nur

so die Variablen identifizieren, die den entscheidenden Unterschied machen.

Forschung zur Wissenschaft und spieltheoretische Analysen zeigen allerdings, dass gute Vorsätze und Empfehlungen oft nicht hinreichend sind, um entsprechende Handlungen hervorzubringen (Auspurg & Hinz 2017; Dawes & Messick 2000; Hinz et al. 2015). Transparente Dokumentationen, das Einstellen von Daten, und das nachhaltige Bereithalten von Materialien für Nachanalysen erfordern Zeit und weitere Ressourcen (wie z.B. Hilfskraftunterstützung, *Server-Space*). Dieser Aufwand könnte ebenso für weitere (Publikations-) Projekte genutzt werden. Zugleich macht sich, wer sich tief in die Karten schauen lässt, besonders angreifbar für Kritik. In einer wissenschaftlichen Umwelt, in der Erfolg und Reputation primär über Publikationsoutput gemessen werden, bedeutet das Wettbewerbsnachteile gegenüber Konkurrent*innen, die weniger in Transparenz und offene Materialien investieren. Aufgrund dieser *Moral Hazards* setzen sich tendenziell nicht-reproduzierbare Aufsätze gegenüber soliden Forschungsbeiträgen durch. Analysen zu dieser Kollektivgutproblematik zeigen, dass in solchen Dilemma-Situationen institutionelle Lösungen gefragt sind: Maßnahmen, welche (a) den Aufwand für die Kooperation (hier: Bereitstellung des Materials) überschaubar halten, und (b) zugleich *verbindliche, einheitliche* Regeln für alle Akteure setzen.

Der *Zeitschrift für Soziologie* ist eine möglichst hohe wissenschaftliche Qualität der publizierten Beiträge von jeher ein hohes Anliegen. Entsprechend wurden in den letzten Jahrzehnten schrittweise Maßnahmen zur Erhöhung der Transparenz quantitativer Beiträge eingeführt. Zunächst im Jahr 2002 mit dem Standard, dass sich Autor*innen quantitativer Artikel vor der finalen Publikationsannahme per Unterschrift zu ausführlichen Datendokumentation (etwa durch deskriptive Statistiken) sowie zur Datenarchivierung verpflichten, um damit ihre empirischen Untersuchungen replizierbar zu machen und auch Fehlerkontrollen zu erleichtern (Editorial der *Zeitschrift für Soziologie* 2002).³ Allerdings zeigten die Erfahrungen, dass diese Selbstverpflichtung allein nicht hinreichend ist. So waren etwa bei einem Lehrforschungsprojekt an der Universität Mannheim im Jahr 2007/2008 unter der Leitung von Josef Brüderl nur 34 der 83 kontaktierten Autorenteams von *Zeitschrift für Soziologie*-Aufsätzen (41 %) bereit, den Bitten nach einer Bereitstellung

² Zugleich sind zugängliche Daten und Auswertungsfiles ein wertvolles Material für die Ausbildung des wissenschaftlichen Nachwuchses. Mit der Möglichkeit der Nachnutzung von Daten lassen sich zudem Drittmittel sowie Zeit und andere Mühen auf Seiten der Beforschten einsparen, die dann für andere Projekte genutzt werden können. Mitunter deswegen machen Forschungsinstitutionen „FAIR data“ (Findable, Accessible, Interoperable, Reusable) immer stärker zur Pflicht.

³ Diese Standards wurden von den damaligen Herausgebenden eingeführt: Andreas Diekmann, Bettina Heintz, Richard Münch, Ilona Ostner, Hartmann Tyrell.

von Daten und Dokumentationen zur Reproduktion der publizierten Analysen nachzukommen. Die Gründe für Nicht-Kooperation waren dabei vielfältig (Brüderl 2013): Einzelne Autor*innen antworten nicht auf Anfragen; andere bedauern einen Datenverlust (womit diese Daten und Auswertungen wohl dauerhaft der Wissenschaft verloren gegangen sind). Daten gehen dabei häufig verloren, weil (Ko-)Autor*innen die Wissenschaft verlassen, Webseiten migrieren und angegebene Links zu Materialien nicht mehr funktionieren; oder einzelne Autor*innen sind letztlich nicht kooperationsbereit.

Alles in allem zeigen diese Erfahrungen, dass die bisherige Praxis nicht gut funktionierte. Daher haben die Herausgebenden der *Zeitschrift für Soziologie* im Jahr 2015 eine weitere Initiative ergriffen, und in Kooperation mit der GESIS einen Replikationsserver für die Archivierung von Daten zu Aufsätzen in der *Zeitschrift für Soziologie* und anderen Zeitschriften entwickelt (<https://www.gesis.org/replikationsserver/home>). Dieser Server ist Teil des seit 2014 bestehenden digitalen Daten-Repositorium *Datorium*, welches die langfristige Sicherung von sozialwissenschaftlichen Daten durch einen nicht-kommerziellen Anbieter garantiert (<https://data.gesis.org/sharing/>).⁴ Für die *Zeitschrift für Soziologie* gilt ab dem Jahr 2016: Autor*innen von Manuskripten, die auf quantitativen Daten basieren, verpflichten sich bei Publikationsannahme alle Materialien, die zum Nachvollzug der Analysen erforderlich sind, auf Datorium oder alternativ einem anderen öffentlichen Replikationsserver zugänglich zu machen. (Neben den Daten und Dokumentationen sind das üblicherweise Aufbereitungs- und Auswertungsfiles, die von den Rohdaten zu den Ergebnissen führen.) Die *Zeitschrift für Soziologie* macht Ausnahmen, die aber der gesonderten Begründung bedürfen. Solche Ausnahmen betreffen vor allem beson-

ders sensible oder prozessproduzierte Daten, bei denen die Autor*innen selbst nur eingeschränkten Zugang haben. Allerdings können auch in solchen Ausnahmefällen in der Regel einzelne Materialien dennoch zugänglich gemacht werden (z. B. Syntaxfiles).

Die Soziale Welt hat im letzten Jahr eine analoge *Data Policy* eingeführt (siehe Jungbauer-Gans et al. 2018), was im Hinblick auf das oben angesprochene Dilemma sehr zu begrüßen ist. Erst wenn alle Zeitschriften mit einheitlichen Regelungen „am gleichen Strang“ ziehen, werden sich die beschriebenen *Moral Hazards* überwinden lassen. In den Nachbardisziplinen der Ökonomie und Politikwissenschaft sind solche Archivierungen daher inzwischen die Regel.⁵

Allerdings funktionieren bekanntlich auch gut überlegte Maßnahmen in der Praxis zuweilen nicht. Im Folgenden nehmen wir daher eine allererste Einschätzung der vorgestellten Archivierungspolitik der *Zeitschrift für Soziologie* vor und fragen: Werden diese Standards für Archivierungen befolgt? Und werden die bereitgestellten Daten auch genutzt, gibt es womöglich unbeabsichtigte Nebenfolgen?

Kooperationsbereitschaft bei der Archivierung: Auswertung des Datenbestands bei Datorium

Die Anforderung der Archivierung quantitativer Analysematerialien gilt verbindlich für Neueinreichungen ab dem Jahr 2016 (die dann ab den späteren 2016er Heften erschienen sind). Die Autor*innen werden nach der Publikationszusage auf Datorium hingewiesen und an die Archivierung erinnert, wenn der Redaktion bis kurz vor Erscheinen des Artikels noch kein zitierfähiger *Digital Object Identifier* (DOI) vorliegt. Dieser wird bei Einreichung der Materialien in Datorium erzeugt und dann idealerweise in die Publikation aufgenommen. Eine weitergehende Überprüfung der Bereitstellung wird bislang nicht vorgenommen.

Gleichwohl zeigt sich eine hohe Kooperationsbereitschaft unter den Autor*innen. Wir verwenden als Datenquelle zunächst eine eigens vorgenommene Kodierung der insgesamt 42 mit quantitativen Daten arbeitenden *Zeit-*

⁴ Insbesondere der damalige Mitherausgeber der *Zeitschrift für Soziologie* Thomas Hinz sowie Tobias Wolbring haben die Entwicklung des Replikationsservers bei Datorium vorangebracht. Forschende haben beim Hochladen von Materialien auf Datorium die Möglichkeit, ihre Einreichung mit Schlagworten aus kontrollierten Vokabularen zu beschreiben. Alle Einreichungen werden durch eine*n GESIS-Mitarbeiter*in u. a. dahingehend überprüft, ob Dateien technisch lesbar sind, ob Daten vollständig sind, und ob sie aus datenschutz- und urheberrechtlicher Perspektive veröffentlicht werden dürfen. Veröffentlichte Materialien sind über verschiedene Suchmaschinen und Kataloge auffindbar – angefangen von der GESIS-Suche bis hin zu Google. Datorium, heute Teil der Plattform SowiDataNet[datorium], ermöglicht Datenbereitstellenden dabei unterschiedliche Zugangs- und Nutzungskategorien auszuwählen und empfiehlt zudem, über die Verwendung einer standardisierten Lizenz wie „Creative Commons“ die Bedingungen der Nachnutzung zu definieren. Mit der Vergabe eines *Digital Object Identifier* (DOI) wird zudem auf die ordnungsgemäße Zitation der Materialien hingewirkt.

⁵ In einigen Zeitschriften in der Ökonomie und Politikwissenschaft geht man inzwischen mit *Data-Editors* sogar noch einen Schritt weiter: Erst wenn von diesem Editor die Reproduzierbarkeit der Ergebnisse nachgeprüft wurde, werden die Manuskripte zur Publikation akzeptiert.

schrift für Soziologie-Artikel, die in den Heften 1/2016 bis 3/2019 erschienen sind. Wir haben in den Artikeln nach Hinweisen auf die Archivierung gesucht und zusätzlich geprüft, ob Daten auf Datorium vorliegen.⁶ Insgesamt war dies bei der Hälfte der Artikel der Fall ($N = 21$). Zusätzlich wurden bei Datorium aber auch noch Daten zu früheren Aufsätzen eingestellt, so dass sich die Zahl der dort verfügbaren Aufsätze auf $N = 26$ beläuft. Die Autor*innen können dabei zwischen unterschiedlichen Kategorien der Nachnutzung wählen. Etwa die Hälfte der 26 Autor*innen ($N = 12$) haben eine „eingeschränkte Nachnutzung“ gewählt, was bedeutet, dass die Nutzung der Daten hier erst nach ihrer expliziten Zustimmung erfolgen kann (wozu die Autor*innen bei Anfragen jeweils von Datorium kontaktiert werden, unter Nennung der Anfragenden und ihrer Forschungsziele). Die übrigen Autor*innen haben ihre Daten „frei“ zur Verfügung gestellt, die Daten können hier entweder direkt von Datorium heruntergeladen werden ($N = 7$ Aufsätze), oder umgehend nach einer vorherigen Registrierung bei Datorium mit Angabe von Kontaktdaten (ebenfalls $N = 7$ Aufsätze).

Wie aus Abbildung 1 ersichtlich wird, ist über die Jahre hinweg die Bereitstellung deutlich angestiegen. Für die ersten Jahre sind vermutlich noch Aufsätze erfasst, deren Erst-Einreichungen vor Einführung der Archivierungs-Policy erfolgt ist (dies ließ sich leider nicht mehr genau zurückverfolgen), was die geringeren Werte für diese Jahre erklären dürfte.⁷ Die in den letzten zwei Jahren (2018, 2019) publizierten Aufsätze sind aber sicher nach Einführung der Policy prozessiert worden. Hier liegen für 75 Prozent der Aufsätze Materialien vor. Bei einem Großteil handelt es sich dabei um die kompletten Replikationsmaterialien (Daten und Syntaxfiles). Damit liegt die Datenzugänglichkeit deutlich über dem, was in den Jahren vor Datorium erreicht wurde (s. Brüderl 2013).

Rücksprachen mit Autor*innen, die bislang wenig oder gar nichts archiviert haben, ergaben, dass die Gründe hierfür oftmals in einer „Verantwortungsdiffusion“ in Autorentteams liegen oder auch bisweilen in Zeitmangel oder kommunikativen Unklarheiten. In fast allen diesen Fällen

wurden die Daten inzwischen nachgereicht. (Wir haben Abbildung 1 danach nicht nochmals aktualisiert, um die Kooperationsbereitschaft ohne explizites „Nachhaken“ abzubilden – inzwischen liegen bei Datorium nochmals deutlich mehr Materialien vor.) Auch andere Rückmeldungen deuten auf ein hohes Verständnis für die Policy der *Zeitschrift für Soziologie* hin. Etwa haben mehrere Autor*innen darauf verwiesen, dass sie selbst in ihren Arbeiten von den Daten von Kolleg*innen profitiert haben. Nur in seltenen Ausnahmefällen lagen unüberwindbare Hürden vor (Datenschutz, etwa bei sensiblen Gesundheitsdaten). In diesen Fällen wurden aber oftmals Syntaxfiles eingereicht.

Als problematisch erwies sich allerdings in wenigen Fällen die Einstellung von Daten mit „eingeschränktem Zugang“, bei dem die Autor*innen vor der Nachnutzung nochmals explizit um Einwilligung gebeten werden müssen. Erfahrungsberichte aus Replikationsseminaren (systematische Daten liegen uns leider nicht vor) haben ergeben, dass hier zuweilen die Autor*innen nicht mehr erreichbar und damit die Daten somit doch nicht zugänglich sind.

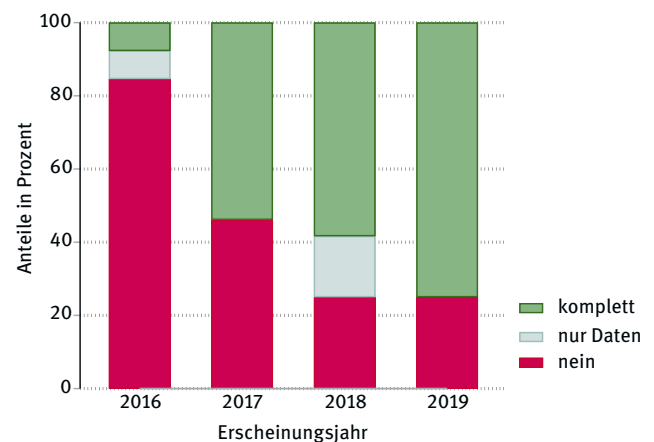


Abb. 1: Datenverfügbarkeit in Prozent nach Erscheinungsjahr (2016–2019)

Anmerkungen: Lediglich Aufsätze mit quantitativer Empirie ($N = 42$). Eigene Recherchen über Aufsätze der *Zeitschrift für Soziologie* und in Datorium. Die Kategorie „komplett“ beinhaltet Daten und Syntaxfiles, und umfasst auch andernorts bestehenden öffentlichen Datenzugang, wie etwa beim Forschungsdatenzentrum des SOEP am DIW Berlin, oder Daten die über die GESIS ohnehin zugänglich sind (z. B. ALLBUS-Daten). Aufgeschlüsselt für die einzelnen Jahre betragen die Fallzahlen jeweils $N = 13$ in 2016 und 2017; $N = 12$ in 2018 und $N = 4$ in 2019.

⁶ Nicht in allen Artikeln wird bislang auf die Bereitstellung hingewiesen. Die Recherche über die Artikel war gleichwohl erforderlich, um auch alternative Repositorien zu Datorium erfassen zu können.

⁷ Die geringere Bereitstellung in den ersten Jahren ist zudem nach persönlicher Kommunikation mit Autor*innen darauf zurückzuführen, dass sie die Archivierung während ihren Projektarbeiten noch nicht vorgesehen hatten, und deshalb Schwierigkeiten hatten, die Dokumentation nachträglich zu erstellen. Gerade dieser Grund verweist nochmals auf die Notwendigkeit einer allgemein stärkeren Verankerung von *Open Science* Praktiken, damit diese schon in den Projektphasen eingeplant werden.

Lohnt der Aufwand?

Ein oft zu vernehmender Einwand gegen die Zweckmäßigkeit einer umfassenden Archivierung ist ein zu hoher Aufwand: Der für die Archivierung erforderliche Aufwand würde nicht in einem angemessenen Verhältnis zum Ertrag stehen, da sich kaum jemand für die Materialien interessieren würde.⁸ Um eine Antwort auf die Frage zu finden, ob die für *Zeitschrift für Soziologie*-Publikationen archivierten Daten bereits Nutzer*innen gefunden haben, ziehen wir eine weitere Datenquelle heran: Prozessproduzierte Daten aus Datorium, die über die Nutzung der dort eingestellten Materialien informieren (andernorts archivierte Materialien können wir in den folgenden Analysen leider nicht berücksichtigen). Insgesamt sind dort (Stand November 2019) Daten zu 26 *Zeitschrift für Soziologie*-Aufsätzen archiviert (davon 25 quantitativ; der weitere Aufsatz arbeitet mit qualitativen Methoden). Die ersten Aufsätze wurden dabei bereits in den späteren 2015er Heften publiziert, die Anfang 2016 in der Druckversion erschienen sind.

Für die Materialien zu diesen 26 Artikeln sind bislang 221 Downloads verzeichnet (Stand 30.11.2019).⁹ Mit einer mittleren Anzahl von 8,5 Downloads pro Artikel stoßen die Materialien auf substanzielles Interesse (Median 3; bei einer Spannweite von 0 bis 33). Wenig überraschend nimmt die Nutzung mit dem Zeitraum zu, in dem Aufsätze bereits publiziert sind (siehe Abbildung 2). Während es für Aufsätze, die in 2019 publiziert wurden, im Mittel lediglich zwei Downloads sind, liegt die mittlere Anzahl für Aufsätze aus den Jahren 2015 und 2016 bereits bei 14 Downloads. Unter den beiden Spitzenreitern mit jeweils 33 Downloads finden sich die Daten einer Primärerhebung (einer postalischen Befragung), die lediglich bei Datorium öffentlich archiviert sind. Auch die Daten des bislang einzigen archivierten qualitativen Artikel sind mit bisher $N = 11$ Downloads auf Interesse gestoßen. Zudem erfahren die Daten des laut *Zeitschrift für Soziologie*-Webseite „meist heruntergeladenen Artikels“ der letzten Jahre von Lengfeld & Dilger (2016) starke Nachfrage ($N = 25$ Downloads; ebenfalls Stand November 2019).

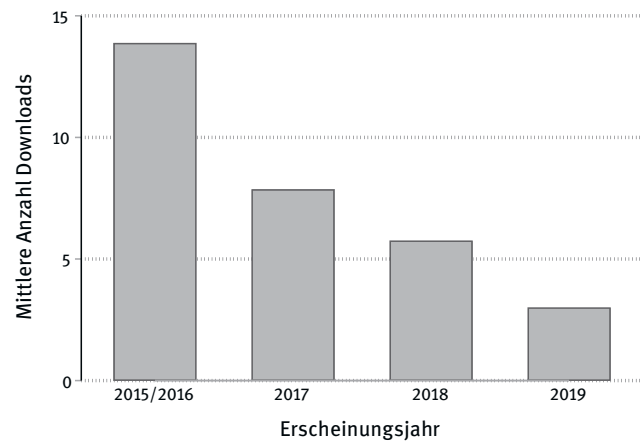


Abb. 2: Mittlere Anzahl von Downloads nach Erscheinungsjahr
 Anmerkungen: $N = 26$ Aufsätze (nach Jahren aufgeschlüsselt: jeweils $N = 8$ Aufsätze in 2015/16 und 2018; $N = 7$ in 2017; und $N = 3$ in 2019).

Ein zentraler Grund für Varianz sind zudem unterschiedliche Zugangskategorien. Bei „eingeschränkter Nutzung“ (Autor*innen müssen explizit um Erlaubnis gefragt werden) ist die mittlere Anzahl an Nutzungen mit $N = 2$ (SD: 2) deutlich geringer als bei den anderen beiden Kategorien. Daten, die eine Registrierung erfordern, wurden im Mittel 17mal heruntergeladen (SD: 9). Für Daten, die völlig frei heruntergeladen werden können, liegt der Mittelwert mit $N = 27$ (SD: 8) Downloads nochmals deutlich darüber.

Abbildung 3 informiert über die Nutzungszwecke. Grundlage ist eine Selbstauskunft, die beim Herunterladen kategorial abgefragt wird. Gut die Hälfte der Downloads ($N = 134$; 61 %) erfolgte im Rahmen von Qualifikationen, insbesondere für Lehrveranstaltungen und BA- und MA-Abschlussarbeiten im Studium.¹⁰ Insgesamt 65 Downloads (29 %) entfallen auf die wissenschaftliche Forschung. Zu welchem Anteil es sich hierbei um Promotionen handelt, ist aus den Daten nicht ersichtlich. Insgesamt spricht das Nutzungsprofil gleichwohl dafür, dass die Datenverfügbarkeit insbesondere dem wissenschaftlichen Nachwuchs zur Qualifizierung zu Gute kommt.

⁸ In dem Zusammenhang sei darauf hingewiesen, dass Forschungsorganisationen wie die DFG inzwischen die Beantragung von Mitteln für die Aufbereitung von Daten und Erhebungsmaterialien für die Nachnutzung explizit vorsehen und anraten.

⁹ Alle Downloads von Materialien zu einem Artikel, die am selben Tag von denselben Nutzer*innen (identifiziert über Session IDs und/oder anonymisierte IP-Adressen) durchgeführt wurden, haben wir zu einem einzigen Download zusammengefasst. Sprich: wenn ein*e Nutzer*in das Codebuch und den Datensatz herunterlädt, zählt dies als ein Download.

¹⁰ Die Kategorie „Qualifikation im Studium“ umfasst mit $N = 126$ (94 %) überwiegend Downloads durch Studierende selbst. In acht Fällen haben Dozent*innen Daten für Lehrveranstaltungen genutzt, wobei in diesen Fällen von einem erweiterten Kreis an Nutzer*innen (Seminarteilnehmende) auszugehen ist.

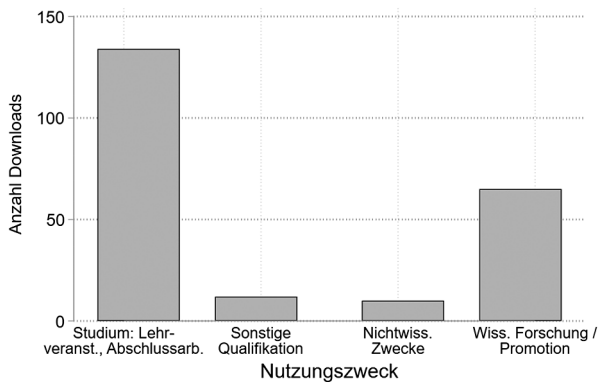


Abb. 3: Angegebener Nutzungszweck

Anmerkungen: $N = 221$ Downloads, verzeichnet für $N = 20$ Artikel (für $N = 6$ Artikel sind bislang keine Downloads verzeichnet).

Unbeabsichtigte Nebenwirkungen?

Zuweilen wird – insbesondere von Herausgebenden oder Verlagen – die Befürchtung geäußert, dass es mit Einführung einer *Data Policy* zu einer Abwanderung von Autor*innen zu alternativen Journals kommen könnte. Eine solche Gefahr sehen wir bei der *Zeitschrift für Soziologie* nicht, weil wir im Gegenteil auf eine Auswahl von möglichst transparenten Beiträgen setzen und dabei auch begründete Ausnahmen zulassen.

Wir können aber auch im Hinblick auf die Bedenken etwas Entwarnung geben: Wir haben den Verlauf der *Impact*-Faktoren und die Verteilung von publizierten Beiträgen auf unterschiedliche Beitragsarten (Artikel mit quantitativem, qualitativen oder theoretischem Schwerpunkt) geprüft. Weder beim *Impact* der *Zeitschrift für Soziologie* noch beim Anteil von quantitativen Beiträgen sind mit Einführung der *Data Policy* auffällige Entwicklungen jenseits der üblichen – teils großen – Schwankungen in diesen Kennwerten erkennbar (betrachtet haben wir alle akzeptierten Manuskripte; siehe die Abbildung A1a und A1b im Online-Anhang zu diesem Beitrag). Diese Zahlen sind mit Vorsicht zu interpretieren, da sie etwa mit anderen Periodeneffekten oder Zeittrends vermischt sein können; und man idealerweise die Einreichungen statt Publikationen betrachten würde, für die uns Daten aber nicht vorliegen. Dennoch sind sie ein weiteres Indiz dafür, dass die meisten mit standardisierten Daten forschenden Kolleg*innen bereit sind, Transparenz mitzutragen. Wie stark die Autor*innen von den Nachnutzungen durch einen stärkeren *Impact* ihrer Beiträge (höhere Zitation) profitieren, ist mit unserem kurzen Beobachtungszeitraum noch nicht zu beantworten. Die intensive Nachnutzung

lässt aber über die Zeit einen Anstieg des Zitations-*Impact* der Aufsätze erwarten, schließlich sind alle nachgenutzten Materialien zu zitieren (für Evidenz zu Aufsätzen in der Ökonomie: McCabe & Mueller-Langer 2019).

Exkurs zur qualitativen Forschung

Bevor wir ein Fazit aus den Analysen ziehen, möchten wir kurz auf die qualitative Forschung eingehen. Denn Leser*innen werden sich fragen, warum die qualitative Forschung bislang aus der *Daten-Policy* der *Zeitschrift für Soziologie* außen vor gelassen ist. Der Grund ist, dass man sich hier allgemein noch nicht auf ähnliche Standards einigen konnte. Das Feld der qualitativen Forschung ist sehr heterogen, die Datenstruktur ist anders (z. B. oftmals mehr sensible Informationen enthaltend), und die Möglichkeit und der Nutzen von Datennachnutzungen und -archivierungen werden aus diesen und weiteren Gründen oft kritischer beurteilt. Auch im Herausgebergremium der *Zeitschrift für Soziologie* konnte entsprechend bisher kein Konsens dahingehend erreicht werden, diese Daten mit einer analogen *Policy* zu versehen oder anderweitige Standards für mehr Transparenz einzuführen. Es ist aber vorgesehen, die Diskussion dazu fortzusetzen. So ist etwa ein *Special-Issue* über Gütekriterien in der qualitativen Forschung geplant. Damit will das Herausgebergremium die auch in anderer Hinsicht noch virulenten Diskussionen zu Gütekriterien für qualitative Methoden aufgreifen (siehe etwa Strübing et al. 2018; Eisewicht & Grenz 2018; Hirschauer et al. 2019).

Wir möchten angesichts des Themas des vorliegenden Beitrags aber noch kurz darauf verweisen, dass inzwischen auch für die qualitative Forschung eine gut ausgebaute Infrastruktur für die Archivierung von Daten und weiteren Materialien zur Verfügung steht, beispielsweise mit dem von der DFG geförderten „Qualiservice“ an der Universität Bremen (<https://www.qualiservice.org/de/ueber.html>). Die Auswertungen von Begutachtungen bei der *Zeitschrift für Soziologie* durch Gunnar Otte (2019) haben gezeigt, dass qualitativ forschende Manuskripte oft an einer zu unklaren Beschreibung der Methodik scheitern, die nicht nur von uns Herausgebenden, sondern insbesondere von einschlägigen Gutachtenden kritisiert wird. Eine möglichst transparente Dokumentation der Methodik hilft mitunter Einwänden zu begegnen, man hätte lediglich selektiv, ohne methodisch kontrollierte Aufarbeitung des Materials zur Untermauerung der eigenen Argumentation aus Interviewpassagen zitiert (Pratt et al. 2019; zur Relevanz von intersubjektiver Nachvollziehbarkeit für

wissenschaftliche Forschung allgemein siehe zudem DFG 2019). Dies ist ein Einwand, der auch in unseren Herausgebersitzungen zuweilen aufgeworfen wird. Will oder kann man nicht die kompletten Daten offenlegen, dann besteht immer noch die Option, einzelne Materialien wie Interview-Leitfäden oder Kodierschemata bereitzustellen. Datorium eignet sich auch dafür, wie einzelne Beispiele bereits zeigen. Zudem können solche Materialien einfach in Online-Anhängen publiziert werden. Weiter bietet die Digitalisierung interessante Optionen, um intersubjektive Nachvollziehbarkeit zu erhöhen, indem etwa längere Zitate verlinkt werden, um damit den Kontext von Interviewpassagen besser zu verdeutlichen (Karcher & Weber 2019). Mit vertieften Einblicken in das Material und seine Aufarbeitung ließe sich gewiss das Verständnis für (insbesondere sehr kreative, nicht-standardisierte) empirische Vorgehensweisen noch erhöhen, und das über Methodengrenzen hinweg.

Fazit: Was folgt aus den Auswertungen?

Insgesamt zeigen unsere Auswertungen, dass sich der Zugang zu quantitativen Daten und Analysematerialien mit Einführung der *Data Policy* bei der *Zeitschrift für Soziologie* erhöht hat und dieses Angebot auch in der wissenschaftlichen *Community* und speziell beim Nachwuchs stark nachgefragt wird.

Gleichwohl sind einige operative Abläufe noch verbesserungswürdig. Wir werden die Hinweise auf der Webseite der *Zeitschrift für Soziologie* präzisieren und gemeinsam mit Datorium überlegen, an welchen anderen Stellen die Kommunikation mit Autor*innen noch optimiert werden kann. Ferner soll die Kategorie des „eingeschränkten Datenzugangs“ für Replikationsmaterialien der *Zeitschrift für Soziologie* abgeschafft werden, da hier immer die Gefahr besteht, dass Autor*innen für die dann erforderliche, separate Einwilligung zu jeder Nachnutzung nicht erreichbar sind. Besonders sensible Daten, die nur eingeschränkt zu nutzen sind, können in besonderen Schutzräumen (*Secure Data Centers*) eingestellt und bearbeitet werden, so dass die *Zeitschrift für Soziologie Data-Policy* dann immer noch umgesetzt werden kann.

Weiter hat sich ergeben, dass oftmals keinerlei Hinweise zu den Daten und Zitationsweisen in den Artikeln zu finden sind. Künftig soll daher jeder empirisch arbeitende Aufsatz mit einer expliziten „Datennotiz“ (*Data Note*) versehen werden, in der die Autor*innen kurz benennen, welche Daten sie verwendet haben (idealerweise

mit exakten Versionsnummern), wo diese verfügbar sind, oder gegebenenfalls, aus welchen Gründen sie nicht bereitgestellt werden können. Diese Datennotizen sollen zugleich über die korrekte Zitierweise informieren, damit die Leistungen der Autor*innen von Nachnutzenden adäquat gewürdigt werden. Mit Umsetzung dieser Datennotizen erfüllt die *Zeitschrift für Soziologie* dann auch als eine der ersten soziologischen Fachzeitschriften die *Top Guidelines* für besonders „offene“ Journals (<https://cos.io/top/>).

Aus den ausgewerteten Kategorien nicht direkt ersichtlich war, inwieweit Nachnutzungen Reproduktionen oder Replikationen dienen. Erwähnt sei daher abschließend, dass die *Zeitschrift für Soziologie* auch für solche Analysen offen ist. Unabhängig von Robustheitsanalysen ergeben sich durch Sekundäranalysen oft substanzielle inhaltliche Erkenntnisse (etwa wenn die Primärauswertungen durch gezielte Ausweitungen von Fällen, Kontextdaten oder Subgruppenanalysen erweitert werden, die dann umfassendere Theorieprüfungen erlauben). Die *Zeitschrift für Soziologie* akzeptiert daher ausdrücklich Reproduktionen sowie Replikationen (erfolgreiche ebenso wie nicht erfolgreiche). Diese unterliegen dem üblichen *Peer-Review* Verfahren. Die Arbeiten haben dabei über reine „push-button-Reproduktionen“ hinauszugehen, sie müssen sich also zumindest durch gut durchdachte Robustheits- und Sensitivitätsanalysen auszeichnen. Zugleich achten wir Herausgebende besonders auf eine wertschätzende Würdigung der Leistungen der Primärautor*innen sowie adäquate Zitation aller bereitgestellten Materialien. Nähere Hinweise zu den möglichen Formaten von Manuskripten finden sich auf unserer Webseite, die in Kürze diesbezüglich auch nochmals aktualisiert wird.

Wir stehen zudem weiterhin mit den Herausgebern benachbarter Journalen im engen Austausch, um sie ebenfalls zur Etablierung von ähnlichen *Data Policies* zu ermuntern. Nur wenn Journal-übergreifend ähnliche Standards gesetzt werden, werden sich die angesprochenen *Moral Hazards* umfassend überwinden lassen. Wir hoffen daher, dass dieses Editorial nochmals dazu anregt, die Standards bei weiteren Journalen zu überdenken.

Was die weitere Öffnung der *Zeitschrift für Soziologie* angeht, sei schließlich noch der ebenfalls zu *Open Science* Praktiken gehörende *Open Access* zu Aufsätzen angesprochen, also der barrierefreie Zugang zu den Publikationen. Mit dem Verlagswechsel der *Zeitschrift für Soziologie* – von Lucius & Lucius zu De Gruyter – wurde ein *Open Access* für alle Artikel spätestens zwei Jahre nach ihrem Erscheinen fortgeführt. Ein Beitrag pro Heft wird zudem von Beginn an freigeschaltet (bevorzugt werden hierfür englischsprachige Artikel ausgewählt, da diese einen besonders hohen Leserkreis erwarten lassen). Autor*innen steht zu-

gleich die Möglichkeit offen, umgehenden *Open Access* zu erwerben, wofür die *Zeitschrift für Soziologie* die Bedingungen und Kosten aktuell neu verhandelt.

Danksagung: Für wertvolle Rückmeldungen zu früheren Versionen danken wir Josef Brüderl, Thomas Hinz, Christian Prinz, Karin Kurz, Herbert Kalthoff sowie den weiteren Herausgebenden der *Zeitschrift für Soziologie*.

Notiz: Dieser Artikel nutzt eigens erhobene prozessproduzierte Daten zur Plattform Datorium sowie zu Publikationen in der *Zeitschrift für Soziologie*. Die kompletten Replikationsfiles (Daten, Stata do-files) sind auf dem Repositorium SowiDataNet|datorium verfügbar: <https://doi.org/10.7802/1992>.



Literatur

- Akademie für Soziologie, 2019: Richtlinien zum Umgang mit Forschungsdaten in der Soziologie. Mannheim. Verfügbar unter (Stand 02.12.2019): https://akademie-soziologie.de/wp-content/uploads/2019/10/Richtlinien-Forschungsdaten-AS_25_09_2019-1.pdf.
- Auspurg, K. & T. Hinz, 2017: Social Dilemmas in Science: Detecting Misconduct and Finding Institutional Solutions. s. 189–214 in: B. Jann & W. Przepiorka (Hrsg.), *Social Dilemmas, Institutions, and the Evolution of Cooperation*. Berlin: De Gruyter Oldenbourg.
- Brüderl, J., 2013: Sind die Sozialwissenschaften wissenschaftlich? Ergebnisse eines Replikationsexperiments. Vortrag am 19.11.2013 auf der Tagung „Rational-Choice-Soziologie“ an der Venice International University.
- Christensen, G. & E. Miguel, 2018: Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature* 56(3): 920–980.
- Christensen, G., J. Freese & E. Miguel, 2019: *Transparent and Reproducible Social Science Research. How To Do Open Science*. Oakland CA: University of California Press.
- Dawes, R.M. & D.M. Messick, 2000: Social Dilemmas. *International Journal of Psychology* 35(2): 111–116.
- Deutsche Forschungsgemeinschaft (DFG), 2019: Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex. Verfügbar unter (Stand 03.12.2019): https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf
- Deutsche Gesellschaft für Soziologie, 2017: Ethikkodex der der Deutschen Gesellschaft für Soziologie (DGS) und des Berufsverbandes Deutscher Soziologinnen und Soziologen (BDS). Verfügbar unter (Stand 2.12.2019): https://soziologie.de/fileadmin/user_upload/dokumente/Ethik-Kodex_2017-06-10.pdf.
- Eisewicht, P. & T. Grenz, 2018: Die (Un)Möglichkeit allgemeiner Gütekriterien in der Qualitativen Forschung – Replik auf den Diskussionsanstoß zu „Gütekriterien qualitativer Forschung“ von Jörg Strübing, Jörg, Stefan Hirschauer, Ruth Ayaß, Uwe Krähnke und Thomas Scheffer. *Zeitschrift für Soziologie* 47(5): 364–373.
- Gelman, A. & E. Loken, 2013: The Garden of Forking Paths: Why Multiple Comparisons can be a Problem, even when there is no „Fishing Expedition“ or „p-Hacking“ and the Research Hypothesis was Posited ahead of Time. Arbeitspapier am Department of Statistics, Columbia University.
- Hinz, T., J. Rössel, K.-O. Maiwald, H. Vollmer & T. Wobbe, 2015: Editorial. *Zeitschrift für Soziologie* 44(1): 2–5.
- Herring, C., 2009: Does Diversity Pay? Race, Gender, and the Business Case for Diversity. *American Sociological Review* 74(2): 208–224.
- Hirschauer, S., J. Strübing, R. Ayaß, U. Krähnke & T. Scheffer, 2019: Von der Notwendigkeit ansatzübergreifender Gütekriterien. Eine Replik auf Paul Eisewicht und Tilo Grenz. *Zeitschrift für Soziologie* 48(1): 92–95.
- Ioannidis, J. P. A., 2005: Why Most Published Research Findings Are False. *PLOS Medicine* 2(8): e124.
- Jasso, G., 1985: Marital Coital Frequency and the Passage of Time: Estimating the Separate Effects of Spouses' Ages and Marital Duration, Birth and Marriage Cohorts, and Period Influences. *American Sociological Review* 50(2): 133–149.
- Jungbauer-Gans, M., C. Kleinert, J. Schupp, M. Trappmann & T. Wolbring, 2018: Bekanntmachung: Regelung zur Archivierung von Daten und Analysefiles quantitativer Studien. *Soziale Welt* 69(3): 351–354.
- Karcher, S. & N. Weber, 2019: Annotation for Transparent Inquiry: Transparent Data and Analysis for Qualitative Research. *IASSIST Quarterly* 43(2): 1–9. <https://doi.org/10.29173/iq959>.
- Lengfeld, H. & C. Dilger, 2016: Kulturelle und ökonomische Bedrohung. Eine Analyse der Ursachen der Parteidentifikation mit der „Alternative für Deutschland“ mit dem Sozio-oekonomischen Panel 2016. *Zeitschrift für Soziologie* 47(3): 181–199.
- McCabe, M.J. & F. Mueller-Langer, 2019: Does Data Disclosure Increase Citations? Empirical Evidence from a Natural Experiment in Leading Economics Journals. JRC Working Papers on Digital Economy 2019–02, Joint Research Centre (Seville site).
- Mellor, D., 2018: The Landscape of Open Data Policies. Center for Open Science: <https://cos.io/blog/landscape-open-data-policies/>.
- Merton, R., 1973 [1942]: The Normative Structure of Science. S. 267–278 in: Ders.: *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.
- Munsch, C., 2015: Her Support, His Support: Money, Masculinity, and Marital Infidelity. *American Sociological Review* 80(3): 469–495.
- Munoz, J. & C. Young, 2018: We Ran 9 Billion Regressions: Eliminating False Positives through Computational Model Robustness. *Sociological Methodology* 48(1): 1–33.
- Otte, G., 2019: Ein Blick in die Black Box des Review-Verfahrens. *Zeitschrift für Soziologie* 48(1): 1–5.
- Pratt, M.G., S. Kaplan & R. Whittington, 2019: Editorial Essay: The Tumult over Transparency: Decoupling Transparency

from Replication in Establishing Trustworthy Qualitative Research. *Administrative Science Quarterly*. <https://doi.org/10.1177/0001839219887663>.

Strübing, J., S. Hirschauer, R. Ayaß, U. Krähnke & T. Scheffer, 2018: Gütekriterien qualitativer Sozialforschung. Ein Diskussionsanstoß. *Zeitschrift für Soziologie* 47(2): 83–100.

Wuttke, A., 2019: Why Too Many Political Science Findings Cannot Be Trusted and What We Can Do About It: A Review of Meta-Scientific Research and a Call for Academic Reform. *Politische Vierteljahresschrift* 60: 1–19.

Zenk-Möltgen, W., E. Akdeniz, A. Katsanidou, V. Naßhoven & E. Balaban, 2018: Factors Influencing the Data Sharing Behavior of Researchers in Sociology and Political Science. *Journal of Documentation* 74(5): 1053–1073.

Zusatzmaterial: Die Onlineversion dieses Artikels bietet Zusatzmaterial (<https://doi.org/10.1515/zfsoz-2020-0001>).