

21 Methodisches Resümee

Der dritte Teil des Resümees fokussiert auf zwei Aspekte der empirischen Untersuchung: auf die Frage der Rolle von quantitativen und qualitativen Daten im Rahmen des Untersuchungsdesigns sowie auf „Zugang“ als Merkmal von Publikationen.

21.1 AUSSAGEKRAFT BIBLIOMETRISCHER DATEN

Bei der Beschreibung des Untersuchungsdesigns¹ wurde darauf hingewiesen, dass mit der Kombination von quantitativen und qualitativen Daten eine wechselseitige Absicherung der Ergebnisse angestrebt wird. Nach Abschluss der Analyse soll das Verhältnis von beiden Datentypen noch einmal abschließend reflektiert werden. Ältere Arbeiten zum Verhältnis eines quantitativen bzw. deduktiv-nomologischen Vorgehens einerseits und einer qualitativ-sinnverstehenden Methodologie andererseits heben die Möglichkeit einer Arbeitsteilung beider Paradigmen in der empirischen Sozialforschung hervor. Dem qualitativen Paradigma käme demnach die Rolle der Phänomenerkundung und Typenbildung, dem quantitativen Paradigma die der statistischen Überprüfung von Zusammenhängen und der Absicherung der Allgemeingültigkeit von Erklärungen zu (etwa: Hoffmann-Riehm 1980. 362).² Der Versuch eines solchen Brückenschlags ist nicht zuletzt auf dem Hintergrund eines längeren und zum Teil erbittert geführten Streits zwischen den Vertretern beider methodologischer Strömungen zu sehen. In der jüngeren Vergangenheit ist zu Recht betont worden, in der Forschungspraxis würden qualitative und quantitative Methoden in verschiedener Weise miteinander kombiniert werden (z. B. Miles und Hubermann 1994: 41f.).

1 Siehe Kap. 8.

2 Siehe als aktuelleres Beispiel auch die tabellarische Gegenüberstellung von qualitativen und quantitativen Untersuchungsansätzen in Häder (2006: 69).

Daher soll gefragt werden: Wie gestalten sich die Verhältnisse der quantitativen und qualitativen Komponenten des Untersuchungsdesigns im Fall der vorliegenden Untersuchung?

Richtet man seinen Blick zunächst auf den Informationsgehalt bibliometrischer Daten, so muss man feststellen, dass er zu einer Erklärung von Open Access kaum geeignet ist. Dies gilt sowohl für die bibliometrischen Daten des Publikationsoutputs der Stichprobe als auch für den Datensatz zur Publikationsinfrastruktur der beiden Fächer. Beide beinhalten im Grunde genommen nur recht dürftige Informationen. Diese eignen sich dazu, bestimmte Ausprägungen und Merkmale von Publikationen oder Publikationsmedien zu beschreiben und miteinander zu vergleichen. Die Analyse der Daten gibt entsprechend Auskunft über den Nutzungsumfang der verschiedenen Spielarten von Open Access, die Merkmale der Journallandschaft eines Fachs und die Präsenz von Autoren mit bestimmten Merkmalen in einzelnen Zeitschriften. Anders formuliert eignen sie sich ausschließlich zur Deskription des Phänomens „Open Access“ und zum Vergleich in unterschiedlichen Kontexten.³ Ebenso wie bei anderen bibliometrischen Untersuchungen ist es auf der Grundlage der hier verwendeten Daten nicht möglich, Ausprägungen des Phänomens und etwaige Unterschiede in einen Kontext zu rücken, der zur Aufklärung des Phänomens beitragen würde – die sozialen Kräfte, die dazu geführt haben, dass Forschungsergebnisse in einer bestimmten Form publiziert wurden, sind nicht von bibliometrischen Merkmale abzulesen. Der Grund für die letztlich große Kontextarmut bibliometrischer Daten ist darin zu sehen, dass sie nicht zum Zweck des Erkenntnisgewinns erhoben wurden und in einem sorgfältig geplanten Prozess der Datenerhebung entstanden sind, sondern Ergebnis einer bestimmten Form der Kommunikation unter Wissenschaftlern bilden, die hier zu Forschungszwecken gewissermaßen ‚nachge-nutzt‘ werden. Verglichen mit selbsterhobenen Daten ist der eingefangene Ausschnitt sozialer Wirklichkeit entsprechend schmal.

Die qualitativen Interviews dienen dazu, diese Leerstelle auszufüllen. Mit ihnen wird aufgeklärt, wie Wissenschaftler in Rückgriff auf bestimmte Formen von Publikationen als Autoren und Rezipienten in das wissenschaftliche Kommunikationssystem inkludiert werden und welche Merkmale die Inklusion aufweist. Sie geben Auskunft darüber, welche Ziele mit der Mobilisierung der Publikationsinfrastruktur verfolgt werden, welche Faktoren darauf Einfluss nehmen und welche Folgeprobleme daraus resultieren. Forschungsstrategisch wurden sie dazu genutzt, das Zusammenwirken der Handlungsroutinen von Autoren und Rezipienten aufzudecken und zu explorieren. Möchte man dieses In-Beziehung-

3 Insbes. Kap. 9.

Setzen eines Phänomens mit Ursachen, bedingenden Faktoren und Folgen für die weitere Operation des Handlungssystems und die Plausibilisierung dieser Zusammenhänge bereits als Erklärung bezeichnen, würde diese auf den qualitativen Interviewdaten basieren. Das eingangs dargestellte, und zum Teil noch in der Methodenliteratur anzutreffende Verständnis der Arbeitsteilung zwischen qualitativer und quantitativer Sozialforschung wäre zumindest der Tendenz nach in sein Gegenteil verkehrt.

21.2 ÜBERLEGUNGEN ZUM MERKMAL „ZUGANG“

Die zweite Überlegung des methodischen Resümees bezieht sich auf das Merkmal „Zugang“. Im Vergleich zu anderen Eigenschaften von Publikationen wie Autorschaft, Erscheinungsort, Publikationstypus, Land und Zitation interessiert es die Bibliometrie erst seit relativ kurzer Zeit.⁴ Abgesehen vom Vergleich der Zitationshäufigkeit von selbstarchivierten gegenüber nicht-selbstarchivierten Publikationen⁵, der Untersuchung der Entwicklung der Repositorien- (z.B. Pinfield et al. 2014) und der Zeitschriftenlandschaft (z.B. Laakso und Björk 2012) liegen erst seit kurzer Zeit Studien vor, die Zugang als Merkmal von Publikationen analysieren (z.B. Gargouri et al. 2012, Wohlgemuth et al. 2017). Das aufkommende Interesse an diesem Merkmal ist gewissermaßen Folge der Digitalisierung. Erst auf der Grundlage der digitalen Publikation sind unterschiedliche Formen des Zugangs entstanden, die das Merkmal in interessanter Weise variieren lassen. Im Folgenden soll kurz auf drei Probleme des Merkmals hingewiesen werden, die im Rahmen dieser Untersuchung deutlich geworden sind: die Zeitabhängigkeit, die Konfundierung mit anderen Dimensionen sowie die Identität einer Publikation.⁶

Eine erste Besonderheit des Merkmals „Zugang“ ist, dass es im Verlauf der Zeit variiert. Die Ausprägung des Merkmals ist dementsprechend abhängig vom

-
- 4 Die gestiegene Bedeutung zeigt sich unter anderem darin, dass das Web of Science seit 2014 Informationen darüber bereithält, ob ein Artikel in einem Gold-Open-Access-Journal publiziert wurde.
 - 5 Siehe Lawrence (2001), Thompson Corporation (2004), Kurtz et al. 2005, Moed (2007, 2012), Davis (2006), Swan (2010), Archambault et al. (2016) und Hersh/Plume (2016) sowie für einen Überblick über die Vielzahl der Studien Wagner (2010) und SPARC (2016).
 - 6 Nicht mehr eingegangen werden soll dagegen auf die klassifikatorischen Schwierigkeiten – siehe hierzu Kap. 6.3.

Zeitpunkt seiner Überprüfung. Eine solche zeitliche Abhängigkeit ist in vielen Bereichen der empirischen Sozialforschung anzutreffen – man denke etwa an das Lebensalter, Haushaltseinkommen und die Präferenzen für eine politische Partei. Für ein publikationsbezogenes Datum ist dies aber eher ungewöhnlich. Eigenschaften der Publikation wie die Anzahl an Autoren, der Erscheinungsort, die Zahl und die Merkmale referenzierter Literatur werden durch die Veröffentlichung fixiert und bleiben in der Regel im Zeitverlauf konstant.⁷ Ein veränderliches Merkmal ist die Zahl der Zitationen, wobei sich, von Ausnahmen abgesehen, die Ausprägung nur in Richtung einer Vergrößerung der Zitationszahl verändern kann. Da zitierende Publikationen über ein Publikationsjahr verfügen, kann der Zeitpunkt der Veränderung festgestellt und der Entwicklungsverlauf der Merkmalsausprägung bestimmt werden. Anders verhält es sich mit Zugang: Er kann sich jederzeit verändern – und zwar sowohl in Richtung einer größeren Offenheit als auch in Richtung stärkerer Restriktionen. Die Gründe dafür sind vielfältig. Autoren können die Manuskripte ihrer Publikation zu jedem Zeitpunkt auf ein Repositorium verbringen und entfernen; subskriptionsfinanzierte Journale können einen Teil oder sämtliche ihrer Publikationen öffentlich zugänglich machen oder von einem Open-Access-Modell zu einem Subskriptionsmodell überwechseln; Bibliotheken und Forschungseinrichtungen können Publikationen retrodigitalisieren. Beispiele für größere Veränderungen, die im Rahmen dieser Untersuchung mit beobachtet wurden, war die Retrodigitalisierung großer Bestände älterer astronomischer Literatur durch *ADS* und die Herstellung eines offenen Zugangs zu einem Teil der Jahrgänge von mehr als 40 Zeitschriften des *Elsevier*-Verlags als Reaktion auf den Boykott „thecostofknowledge“.⁸ Bezüglich der Untersuchung des Merkmals „Zugang“ kommt erschwerend hinzu, dass der Zeitpunkt der Veränderung von Zugangsmodalitäten meist nicht dokumentiert ist.⁹ Diese hat mindestens zwei Konsequenzen. Durch die Variabilität des

-
- 7 Ausnahmen bilden etwa Rücknahmen publizierter Artikel (Retractions), die auf wissenschaftliches Fehlverhalten oder Irrtümer zurückgehen. Diese sind aber vergleichsweise selten (siehe hierzu auch die Plattform *Retraction Watch*: <http://retractionwatch.com/> Zugriff am 18. April 2017).
- 8 Aber auch Veränderungen in Richtung einer Verringerung des Zugangs wurden beobachtet. Ein Beispiel stellt die weiter oben für die Mathematik beschriebene Selbstarchivierungspraxis dar: Publikationen werden z.T. nach der Fertigstellung des Manuskripts auf einem Repositorium abgelegt und bei Erscheinen am originären Publikationsort wieder gelöscht.
- 9 Eine Ausnahme stellen hier manche Repositorien wie das *arXiv* dar, die den Zeitpunkt der Selbstarchivierung mit einem Zeitstempel versehen.

Merkmals sind empirische Ergebnisse in einem hohen Maße zeitabhängig, können also erstens nur Geltung für den Zeitpunkt ihrer Erhebung beanspruchen. Will man Aussagen über die derzeitige Zugänglichkeit von Publikationen treffen, deren Zugang zu einem früheren Zeitpunkt bereits einmal erhoben wurde, muss eine erneute Erhebung des Merkmals stattfinden. Zweitens verbietet es die Variabilität des Merkmals, die Verteilung des Merkmals „Zugang“ aus Publikationen verschiedener Jahrgänge als Entwicklung zu interpretieren. Festgestellt werden kann lediglich, in welchem Umfang für welches Publikationsjahr welche Art von Zugänglichkeit besteht. Um Aussagen über Entwicklungen treffen zu können, wäre es notwendig, eine Zugangsprüfung zu mehreren Zeitpunkten vorzunehmen.

Eine zweite Eigenschaft von „Zugang“ ist seine Konfundierung mit anderen Merkmalen. Hierdurch ergeben sich unterschiedliche Möglichkeiten, die Ausprägungen von „Zugang“ zu definieren. Deutlich wird dieser Punkt, wenn man von der einfachsten Möglichkeit ausgeht, Zugang dichotomisiert und zwischen einem leserseitigen freien Zugang und einem (wie auch immer gearteten) beschränkten Zugang unterscheidet. Bei der Anwendung dieser Unterscheidung irritieren regelmäßig Publikationen, von denen mehrere Versionen existieren. Um hier eine eindeutige Zuordnung zu einer Merkmalsausprägung vornehmen zu können, wäre es im Grunde notwendig, das Merkmal nicht auf die Publikation, sondern jeweils auf die *Version* einer Publikation zu beziehen. Bei genauerer Betrachtung sind zwei Fälle zu unterscheiden: Zum einen kann es sich bei den beiden Versionen um eine gedruckte und eine digitale Version handeln. Die gedruckte Version ist dabei zwangsläufig zugangsbeschränkt, weil es ihr aufgrund der medialen Basistechnologie an der Voraussetzung für einen freien Zugang fehlt. Anders bei der digitalen Version: Hier wäre es möglich, dass sie frei zugänglich oder zugangsbeschränkt ist. Dieser erste Fall macht deutlich, dass das Merkmal „Zugang“ partiell mit der „medialen Basistechnologie“ (gedruckt oder digital) konfundiert ist. Zum anderen können beide Versionen in einem digitalen Format vorliegen und in diesem Fall können dann keine¹⁰, eine oder beide Versionen frei zugänglich sein. Unproblematisch ist nur die erste Möglichkeit, die eindeutig als zugangsbeschränkt zu klassifizieren ist, während die zweite zu der Frage führt, ob man bei der Zuordnung nun der frei zugänglichen oder der zugangsbeschränkten Version den Vorzug geben soll. Die dritte Möglichkeit lässt zwar eine eindeutige Zuordnung zu, führt aber unmittelbar zur Frage nach dem

10 Man denke hier an den Fall einer zugangsbeschränkten Publikation, die durch einen Verlag vertrieben werden und zudem in kostenpflichtigen Volltext-Archiven (z.B. *JSTOR*) bereitgehalten werden.

Verhältnis der beiden frei zugänglichen Versionen zueinander und macht die Einführung weiterer Unterscheidungen notwendig. Die Konfundierung des Merkmals „Zugang“ mit „medialer Basistechnologie“ („freier Zugang hat ein digitales Format zur Voraussetzung“) und mit „freier Zugangsart“ („eine Publikation kann auf unterschiedlichem Weg frei zugänglich sein“) machen es notwendig, im Rahmen des Forschungsprozesses Entscheidungen darüber zu treffen, wie die Ausprägung des Merkmals definiert wird. Eine abschließende Verständigung auf Arbeitsdefinitionen innerhalb der bibliometrischen Community steht hier noch aus.¹¹

Ein dritter Aspekt, der während der Zugangsprüfung aufgefallen ist, ist die Frage nach der *Identität einer Publikation*. Während diese in vor-digitalen Zeiten vergleichsweise unproblematisch mit Verweis auf den im originären Publikationsmedium erschienenen Text zu klären war, ist bei der digitalen Publikation das Vorliegen mehrerer Versionen fast die Regel.¹² Probleme bei der Bestimmung der Einheit „Publikation“ treten dann auf, wenn die verschiedenen Versionen in einem mehr oder weniger großen Umfang voneinander abweichen. Unproblematisch ist hier noch der Fall einer selbstarchivierten Publikation in der nicht gesetzten Autorenversion, die sich durch das Fehlen des Verlagslayouts und der Paginierung auszeichnen. Wenn aber, wie es im Rahmen dieser Untersuchung regelmäßig der Fall war, zwei Dokumente an so zentralen Punkten wie dem Titel, der Liste und der Reihenfolge der Autoren, dem Abstract und der Struktur des Volltextes erhebliche Differenzen aufweisen, stellt sich die Frage, wie viel Übereinstimmung von Text und Metadaten gegeben sein muss, um noch von *einer* Publikation sprechen zu können. Datenbanken wie *ADS* und Suchmaschinen wie *Google Scholar*, die Publikationen und E-Prints einander zuordnen, helfen zwar praktisch bei der Zuordnung verschiedener Versionen einer Publikation, können aber kaum darüber hinwegtäuschen, dass durch die digitalen Publikationsmöglichkeiten und insbesondere die Möglichkeiten der Selbstarchivierung die Identität der Publikation unscharf geworden ist. Verlässt man sich auf die Zuordnungsleistung dieser Datenbanken, stellt sich die Frage nach deren Zuverlässigkeit.

-
- 11 Angesichts der Bestrebungen, einen Monitor der im Open Access verfügbaren Publikationsanteile einzurichten (z.B. Abgeordnetenhaus Berlin 2015; BMBF 2016; European Commission 2016), ist es wahrscheinlich, dass die Merkmalsausprägungen künftig noch deutlich stärker aufgliedert und im Bereich von Open Access zwischen Hybrid, Moving Wall und True Open Access unterschieden wird (siehe Kap. 6.2).
 - 12 Ein Extremfall bildete bei der empirischen Zugänglichkeitsprüfung eine Publikation, für die *Google Scholar* 21 Versionen nachwies.