

*Cornelia Schendzielorz, Martin Reinhart*

## Die Regierung der Wissenschaft im Peer Review

### Zusammenfassung

Dieser Beitrag widmet sich der fallvergleichenden Analyse des unübersichtlichen Evaluierungsphänomens *Peer Review*. Ausgehend von der Vielfalt des Peer Review als Gegenstand der Soziologie des Bewertens wird eine Untersuchungsperspektive entwickelt, die Peer Review als Entscheidungsmechanismus über knappe Ressourcen in seiner Verfahrenförmigkeit und Steuerungsfunktion einer empirisch vergleichenden Verfahrensanalyse zugänglich macht. Der Forschungsansatz kombiniert die praxeologischen (Stefan Hirschauer) und funktionalistischen Verfahrensanalysen (Niklas Luhmann) mit konzeptionellen Ansätzen aus der Soziologie des Bewertens und ermöglicht es, relevante Parameter zur Differenzierung und Typologisierung von Peer Review in der Wissenschaft zu bestimmen. Anhand dessen werden spezifische Tätigkeiten herausgearbeitet, die als elementare Bestandteile von Peer-Review-Verfahren identifizierbar sind. Unterschiedliche Peer-Review-Prozesse lassen sich damit anhand ihres Komplexitätsgrades bestimmen und im Spannungsverhältnis von Selbst- und Fremdsteuerung als Evaluationspolitik verorten.

*Schlagwörter:* Peer Review, Legitimation durch Verfahren, Soziologie der Bewertung, Evaluationspolitik, Wissenschaft

### Abstract

#### *Governing Science Through Peer Review*

This article deals with the comparative analysis of peer review as a complex evaluation phenomenon. Starting from the variety of peer review as an object for the sociology of evaluation, we develop a research perspective that makes peer review in its procedural form and control function accessible to an empirical comparative procedural analysis comprehending it as a decision-making mechanism for scarce resources. Our research approach combines the praxeological (Stefan Hirschauer) and functionalist analyses of processes (Niklas Luhmann) with the conceptual approaches from the sociology of evaluation and enables to determine relevant parameters for differentiating and typologising peer review in science. This allows the identification of specific activities that can be identified as elementary components of peer review procedures. Differing peer review processes can thus be determined according to their degree of complexity and positioned as evaluation policy between internal and external forms of control.

*Keywords:* Peer review, legitimation by procedures, sociology of evaluation, evaluation policy, science

## 1 Einleitung

Dieser Beitrag richtet das Augenmerk auf die „varieties of evaluation“, die aktuell in der Wissenschaft vorzufinden sind. Gegenstand der Analyse sind die Evaluierungsarrangements des wissenschaftlichen Peer Reviews unter Berücksichtigung der wissen-

schaftspolitischen Dimension dieser Evaluationsverfahren, die sich u. a. darin zeigt, dass sie über die Verteilung knapper Güter eine Steuerungsfunktion ausüben. Das Peer Review ist nicht nur das zentrale Verfahren wissenschaftlicher Wertbestimmung und damit „one of the fundamental conditions of possibility of academic knowledge and the construction of its value“ (Biagioli, 2002, p. 11). Es ist zudem gleichermaßen bemerkenswert stabil, wie es in seinen konkreten Ausprägungen in spezifischen Prozeduren erstaunlich divers erscheint. Es ist einerseits vielseitig einsetzbar und lässt sich andererseits an den jeweiligen Bedarf anpassen. So sind sowohl Entscheidungsprozesse über Publikationsplatz als auch über Fördermittel oder Stellen mittels Peer Review organisierbar, weil sich dieses an die jeweiligen Erfordernisse anpassen lässt, bspw. durch verschiedene Formen der Verblindung, Arbeitsteilung und Rollentrennung oder durch mehrstufige Entscheidungsprozesse (Biagioli, 2002; Baldwin, 2018). Ähnlich der Evaluierungslandschaft anderer Felder ist es einerseits von Homogenisierungs- andererseits von Heterogenisierungstendenzen geprägt, die nicht zuletzt mit den diversen organisationalen Rahmenbedingungen – Verlag, Zeitschrift, Förderorganisation, Hochschule – zusammenhängen dürften. Somit konstatieren wir angesichts seiner Erfolgsgeschichte als Evaluationsform in der Wissenschaft eine „Metastabilität des Peer Reviews“ (Reinhart, 2012, S. 193), die empirisch mit einer Vielfalt von Peer-Review-Formaten einher geht.

Diese Vielfalt spiegelt sich auch in der Peer-Review-Forschung wider, die sich in großen Teilen dem Zeitschriften-Peer-Review widmet und zur quantitativen Analyse und Prüfung der Reliabilität, Fairness und Validität der Verfahren neigt (Weller, 2001; Hirschauer, 2005; Reinhart, 2012; Neidhardt, 2016). Diese Ausrichtung ist maßgeblich durch die nur selektiv und selten im synchronen oder diachronen Vergleich zugänglichen Daten bedingt. Der mitunter von den Organisationen, die Peer Review nutzen und mitorganisieren - konkret der Wissenschaftsverlage, Förderorganisationen und Hochschulen - begrenzte oder gar verweigerte Zugang zu Daten, die im Laufe jener Evaluationsarrangements generiert werden, stellt seinerseits einen evaluationspolitisch relevanten Befund dar. Dennoch ist angesichts der Bedeutsamkeit des Peer Reviews in und für die Wissenschaft ausgehend von den vielfältigen Fallstudien mit entsprechend fallspezifischen Ergebnissen eine Systematisierung mit Blick auf eine mögliche Theoretisierung wünschenswert. Vor diesem Hintergrund folgt dieser Beitrag dem Desiderat einer fallvergleichenden Analyse des unübersichtlichen Evaluierungsphänomens Peer Review. Dabei wird auf der Basis von Sekundäranalysen umfangreicher empirischer Studien zu Forschungsförderprogrammen ein konzeptionelles Argument zur Differenzierung und Typologisierung von Peer-Review-Prozessen entwickelt.

Dabei ist zu berücksichtigen, dass Peer Review neben der Qualitätssicherung durch wissenschaftliche Expertenurteile auch ein wesentlicher Bestandteil der Evaluierungsarrangements ist, mit denen Wissenschaft beansprucht, sich selber zu steuern. Die politische Dimension jener Selbststeuerung manifestiert sich darin, dass zur Rechtfertigung einer Autonomie und Selbstverwaltung der Wissenschaft auf Peer-Review-Verfahren zurückgegriffen wird (Wilholt, 2012). Gleichwohl wird Peer Review beispielweise im Zuge themenbezogener, kompetitiver Vergabe von Förder- und Forschungsgeldern ebenso zur Fremdsteuerung genutzt. Die Funktionalität und Stabilität des Peer Review gründet offenbar gerade darin, dass es ermöglicht zwischen Selbst- und Fremdsteuerung zu vermitteln und hier eine Art Scharnier bildet. Dies hat in der Forschungsliteratur überwiegend dazu geführt, diese Steuerungskonstellation unter dem Begriff der

Governance zu fassen (Schimank, 2016). Im Gegensatz dazu laufen unsere Theoretisierungsbestrebungen im vorliegenden Beitrag auf den Begriff der Regierung in einem allgemein politischen Sinne (Rosanvallon, 2018) hinaus. Im Folgenden werden wir verschiedene Fallbeispiele und Anwendungskontexte von Peer Review in einer vergleichenden Verfahrensanalyse zusammenführen. Dabei gilt es, die Anpassungsfähigkeit des Peer Review an vielfältige Evaluierungszwecke mit Blick auf jene Scharnierfunktion aufzuschlüsseln. Ziel der Analyse ist es, relevante Parameter zur Differenzierung und Typologisierung von Peer Review in der Wissenschaft herauszuarbeiten, ohne dabei die Einheit des Phänomens Peer Review aus den Augen zu verlieren. Der Beitrag nimmt Peer Review somit als einen Aspekt der Evaluationspolitik von Wissenschaft in den Blick, indem er spezifische Peer-Review-Formate als Bestandteile eines umfassenderen Systems der Selbst- und Fremdsteuerung und damit der Regierung von Wissenschaft untersucht.

## 2 Die Vielfalt des Peer Review als Gegenstand der Soziologie des Bewertens: Bestandsaufnahme und konzeptionelle Überlegungen

Peer Review findet in der Wissenschaft an zentralen Stellen Anwendung und leistet dort Qualitätssicherung über Bewertungsprozesse, insbesondere in Prozeduren zur Manuskriptbegutachtung bei Zeitschriften und Verlagen, zur Antragbegutachtung bei der Einwerbung von Drittmitteln und in der Begutachtung von Personen mit Blick auf ihre Eignung für langfristige Anstellungen, bspw. bei der Berufung von Professor\*innen. In allen drei genannten Fällen, deren Peer Review auch innerhalb eines Anwendungskontextes und desselben organisationalen Rahmens sehr heterogen sein kann, geht es erstens um die Frage der Verteilung knapper Güter, zweitens sind die Peer-Review-Prozesse verfahrensförmig organisiert und drittens erfüllen sie eine Steuerungsfunktion. Angesichts der bereits viel beforschten Frage, wie die knappen Ressourcen verteilt werden, bzw. ob sie fair, wissenschaftlich oder gesellschaftlich sinnvoll verteilt werden sollten<sup>1</sup>, konzentrieren wir uns zunächst auf die Verfahrensförmigkeit des Evaluationsarrangements Peer Review und erörtern anschließend seine Steuerungsfunktion als ein Instrument der Regierung von Wissenschaft. Dabei gilt es, die Verfahren als durch spezifische Praktiken gekennzeichnete in den Blick zu nehmen und zu fragen, in welchem Maße eine Legitimation durch Verfahren erfolgt und inwieweit diese für die Steuerungsfunktion maßgeblich ist. Im Folgenden schließen wir daher an Stefan Hirschauers Verständnis an, demnach Peer Review als ein sozialer (und primär kommunikativer) Prozess begriffen und als solcher in seiner Praxeologie analysiert werden kann (Hirschauer, 2005). Auf dieser Basis entwickeln wir unseren Analyseansatz in kritischer Auseinandersetzung mit der Soziologie der Bewertung (Lamont, 2012; Heintz, 2018) und in Anlehnung an Niklas Luhmanns Begriff des Verfahrens als einer Prozedur, durch die Legitimation erzeugt wird, weiter (Luhmann, 1983).

Eine praxeologische Betrachtung von Peer-Review-Prozessen kann auf Stefan Hirschauers Analyse des Zeitschriften-Peer-Reviews zurückgreifen: Peer Review ist darauf ausgerichtet, die Qualität der publizierten Fachbeiträge zu sichern, die Kommunikationsangebote zu optimieren und dadurch die Lesezeit einer Disziplin zu kalibrieren

(Hirschauer, 2019, S. 227). Die Bewertung erfolgt im Prozess der Urteilsfindung, im Zuge derer diverse Lesemodi von Reviewer\*innen und Editor\*innen – tendenziell inhaltsfokussiert bei Fachgutachter\*innen, tendenziell form- bzw. stilfokussierte bei Herausgeber\*innen – diskutiert werden. Diese Erörterungen ereignen sich beispielsweise in Herausgebersitzungen oder verschriftlicht in Form von Gutachten und der auf diese eingehenden Überarbeitungen der Texte durch die Autor\*innen (Hirschauer, 2005, S. 59-80). In diesem inhaltlichen Austausch zeigt sich bereits die sequenzialisierte Struktur der Prozesse, die sich auch im jeweiligen Begutachtungsverfahren manifestiert, in das sowohl schriftliche als auch mündliche Kommunikationsakte eingehen (Hirschauer, 2019, S. 278-281; 2015, S. 38).<sup>2</sup> Die Charakteristiken der Prozessualität, der Sequenzialität, der Verfahrensförmigkeit und die Mischung von Schriftlichkeit und Mündlichkeit gelten ebenso für Peer-Review-Formate zur Vergabe von Forschungsgeldern und Stellenbesetzungen. Auch in diesen Fällen handelt es sich um komplexe Bewertungsprozesse. Sie sind in mehrstufigen Verfahren organisiert, die zumindest eine grobe sequenzielle Anordnung der jeweiligen Bewertungspraktiken vorbahnen. Dabei werden ebenfalls meist schriftliche Gutachten eingeholt und anschließend Einschätzungen und Lesarten in Gutachtersitzungen bei Förderorganisationen (Olbrecht, 2014) oder in Berufungskommissionen (Kleimann & Klawitter, 2017a) kontrovers diskutiert und erörtert.

Wie ist nun der Bewertungsprozess zu konzeptionalisieren? Im Rahmen der Soziologie der Bewertung findet sich der Vorschlag, Bewertungskonstellationen als Ausgangspunkt zu verwenden, die als Triade von Bewertetem, Bewertendem und Publikum konstituiert sind (Meier, Peetz & Waibel, 2016). Den Autoren zufolge lassen sich diese Konstellationen darüber hinaus vor dem Hintergrund transituativ gültiger Regeln der Bewertung und technologischer Infrastrukturen fassen. Dieser Ansatz zielt darauf ab, sich von einem akteurszentrierten Zugriff der Analyse von Bewertungen zu lösen. Allerdings führt diese Triade die Unterscheidung zwischen Subjekt und Objekt der Bewertung, und somit der Differenzierung zwischen einem aktiv agierenden und passiv bezeichneten Element fort. Zudem behält sie mit dem Begriff des Publikums auch in der Nomenklatur Spuren einer akteurszentrierten Perspektive bei. Als möglicher Ausgangspunkt oder Fixpunkt verbliebe dann einzig der fallspezifisch zu beschreibende Gegenstand der Bewertung. Einen solchen als (ex ante) klar bestimmbar anzunehmen, kollidiert allerdings mit einem praxeologischen resp. pragmatischen Verständnis von (Be-)Wertung. Demnach ist die Ungewissheit Ausgangspunkt und Anlass der valueierenden Untersuchung, die der Bewertung zu Grunde liegt (Dewey, 1939; Potthast, 2017; Bogusz, 2018). Die Bewertungskonstellation lässt sich folglich mit der vorgeschlagenen Triade nicht hinreichend trennscharf erfassen und begrifflich bestimmen, um sie als heuristisches Analyseinstrument für eine vergleichende Verfahrensanalyse nutzbar zu machen.

Demgegenüber identifizieren Anne Krüger und Martin Reinhart (2016; 2017) fünf Bausteine der Bewertung: Bewertungspraktiken, Bewertungsstrukturen, Bewertungsinfrastrukturen, Bewertungssituationen und die Reflexivität von Bewertung. Außerdem plädieren sie innerhalb der Bewertungspraktiken für eine Differenzierung von Wertzuschreibung und Wertabwägung. Diese Unterscheidung aufgreifend ließe sich zur Charakterisierung der Verfahren und ihrer Sequenzialität fragen: Wo, d. h. an welchem Punkt im Bewertungsprozess, erfolgt auf welcher Ebene von wem und wie, d. h. in welcher Form, die *Wertzuschreibung* und die *Wertabwägung*? Das Potential der Unter-

scheidung für eine vergleichende Verfahrensanalyse ist jedoch auch damit begrenzt, dass die Separierung beider Praktiken fall- und verfahrensspezifisch stark variiert und somit schwer systematisierbar ist. Der konkrete Vollzug der Wertzuschreibung und ihrer Hervorbringung ist zudem bei vielen Verfahren in der empirischen Untersuchung schwer zugänglich; denn diese wird häufig als Werteinschätzung individuell erarbeitet und erst anschließend als bereits gefasste Wertzuschreibung erläutert und expliziert. Somit können im Zuge der Peer-Review-Verfahren häufig nur Bestätigungen, Infragestellungen, kontroverse Diskussionen, Modifikationen, Plausibilisierungen und Rechtfertigungen von Wertzuschreibungen beobachtet werden. Damit geht einher, dass die Thematisierung der erfolgten Wertzuschreibungen in Peer-Review-Verfahren häufig unter Legitimationsdruck erfolgen und dementsprechend rhetorisch rationalisiert werden. Dabei werden die affektbehafteten Elemente der Wertzuschreibung häufig euphemisiert, um ihre Akzeptanz als scheinbar rein in der Sache liegende Argumente nicht zu gefährden. Damit weist die Differenzierung von Wertzuschreibung und Wertabwägung in eine ähnliche Richtung wie Hirschauers Verständnis eines relationierten Fachurteils. Überdies zeichnen sich Konvergenzen ab, zwischen dem als „Reflexivität der Bewertung“ gefassten Hinterfragen der eigenen Wertzuschreibung und Wertabwägung und dem, was Hirschauer als „Urteilen unter Beobachtung“ beschreibt (Hirschauer, 2019). Letzteres impliziert, dass Aufmerksamkeit und Sachlichkeit derart intensiviert werden, dass sie zur Steigerung von Meinungsverschiedenheit führen, die dann – zugespitzt formuliert – im „Beurteilungspanoptismus“ (Hirschauer, 2019, S. 287, 290) mündet und reflexiv relationiert werden muss.

Inwiefern trägt nun die Verfahrensförmigkeit, also die prozesshafte, sequenziell formalisierte Organisation der Praktiken, die im Peer-Review-Prozess zusammenkommen zur Legitimation jener Qualitätsbewertung und (Selbst-)Steuerung bei? Wir gehen davon aus, dass eine gelingende Legitimation von Verfahren zur Qualitätssicherung über den Aspekt der personalen Teilnahme oder der Repräsentation entscheidend vermittelt ist (Luhmann, 1983). In dieser Perspektive hat ein Verfahren kein Wahrheitskriterium, fördert die Richtigkeit der Entscheidung, ermöglicht und kanalisiert Kommunikation und garantiert, dass Entscheidungen getroffen werden (Luhmann, 1983, S. 12-13). Demzufolge lässt sich die Verfahrensgestaltung von Prozessen der Qualitätsentwicklung und Bestimmung von Qualität anhand der folgenden fünf Aspekte bestimmen (Luhmann, 1983, S. 40-49; Reinhart, 2012, S. 134-145):

- Offenheit des Verfahrens,
- Verlässlichkeit, mit der es eine Entscheidung herbeiführt,
- Stufenförmigkeit, d. h. die Unterteilung in Sequenzen und Verfahrensschritte,
- Grad der Arbeitsteilung, d. h. die tätigkeitsbezogene, inhaltliche oder funktionale Aufspaltung der zu verrichtenden Aufgaben und schließlich
- Rollendifferenzierung, also die (personenbezogene) Differenzierung von funktionalen Zuständigkeiten, ggf. nach Professionalisierungsgrad, beruflicher Position, organisationalem Status und Repräsentationszwecken.

Um diese konzeptionellen Überlegungen im Sinne einer Heuristik zur vergleichenden Verfahrensanalyse von Bewertungsprozessen in der Wissenschaft fruchtbar zu machen, bilanzieren wir die Divergenzen und Konvergenzen der herangezogenen Literatur. Demzufolge lassen sich Peer-Review-Prozesse in der Wissenschaft als Bewertungsverfahren fassen, in welchen dem bewerteten Objekt ein spezifischer Wert zugeschrieben

und dieser dann zu anderen Bewertungen ins Verhältnis gesetzt und abgewogen wird (Krüger & Reinhart, 2016). Auf diese Weise wird in allen Fällen eine relationale Bewertung hervorgebracht, die im Zuge der Entscheidungsfindung darüber, ob das Manuskript publikationswürdig, das Forschungsvorhaben förderungsfähig oder der Stellenanwärter gewinnungswürdig ist, konsolidiert wird. Peer Review in der Wissenschaft ist somit ein durch das Verfahren und das Fachurteil legitimerter Bewertungsprozess, der Steuerungsfunktionen erfüllt.

Wir plädieren für einen Ansatz, der vom Verfahren aus denkt und somit ermöglicht, im Verfahren zwischen Bewertungspraktiken, Infrastrukturierung, Regeln und Kriterien zu differenzieren. Das Verfahren bietet dabei eine transssituative Analyseeinheit, die es erlaubt, diverse Elemente des Verfahrens zu erfassen, in der empirischen Analyse deren Spezifika aufzuzeigen und in exemplarischen Untersuchungen der Tatsache Rechnung zu tragen, dass distinkte Schritte des Verfahrens in konkreten Situationen vollzogen werden und verankert sind. Somit können Infrastrukturen (Star, 1999; Star & Ruhleder, 1996; Star & Bowker, 2002; Niewöhner, 2014) als Manifestationen spezifischer Verfahrensschritte oder manchmal des gesamten Verfahrens gelesen werden. Derart können sie ggf. auch als „Framing“ im Goffmanschen Sinne (Goffman, 1980) fungieren. Zudem ermöglicht der Fokus auf das Verfahren als eine transssituative Analyseeinheit auch mit Blick auf das Ziel der Bewertungsverfahren deutlicher zu differenzieren. Daraus folgt die Unterscheidung ziel- und zweckspezifischer Regeln, welche die Bewertungsprozesse im Sinne von teils expliziten, teils impliziten Spielregeln kennzeichnen. Dementsprechend sind Bewertungspraktiken als multivalent zu denken, insofern in ihnen Regeln sowohl reproduziert, aktualisiert, als auch modifiziert und teilweise konstituiert und infrastrukturisiert werden. Infrastrukturierung kommt deshalb eine gesonderte Bedeutung zu, weil sie als „materiell-semiotische Praxis“ Beziehungen herstellt und unterbindet, wodurch soziale Praxis als Verflechtung von Akteuren, Technologien und Wertordnungen artikuliert wird (Niewöhner, 2015, p. 123).

Angesicht der dargelegten konzeptionellen Unschärfe insbesondere des Publikumsbegriffs favorisieren wir, die diesem Begriff inhärenten Fragen der Involvierung, Partizipation und Repräsentation in Bewertungsprozessen alternativ mit Erving Goffmans Konzept des Teilnahmestatus zu fassen. Dieses dient dazu, genauer zu unterscheiden, wie jemand einer Situation beiwohnt: Aktiv, akzeptiert, passiv, ausgeschlossen, geduldet, etc. Derart können auch in komplexen Peer-Review-Verfahren die Modi der situativen Beteiligung differenzierter gefasst werden und beispielweise die recht groben Begriffe des Publikums, Zuschauers oder Zuhörers dekonstruiert werden (Goffman, 1986, S. 40-42; 2005, S. 42 ff.; Goffman, 1986, S. 41, Fn. 24; 2005, S. 51, 54, 58, 61, 68). Goffman weiterdenkend, der den Begriff zur Analyse sprachlicher Äußerungen in physischer Kopräsenz nutzt, schlagen wir vor, den Teilnahmestatus weniger in Bezug auf Situationen, sondern mehr in Bezug auf einzelne Verfahrensschritte zu denken, indem man die Aufgaben und Praktiken der Beteiligten herausarbeitet. In diesem Kontext wird auch Goffmans Rahmenbegriff nicht so sehr hinsichtlich der Situationen relevant gemacht, sondern ist als „Framing“ bedeutsam, das sich in der Verfahrensgestaltung vermutlich eher in infrastrukturierenden Regeln und Abläufen manifestiert, welche die Handlungspraktiken der Beteiligten vorbahnen.

### 3 Methodologie

Ausgehend von diesen konzeptuellen Überlegungen zielen wir auf eine vergleichende Betrachtung von Peer-Review-Verfahren. Um relevante Parameter zur Differenzierung und Typisierung von Bewertungspraktiken zu identifizieren, schlagen wir folgende Frageheuristik vor:

- Erstens: Zu welchem *Ziel und Zweck* wird die Bewertung vorgenommen? Die Frage richtet sich auf das verfahrensspezifische Ziel, das bspw. in einer funktionalen Zweckbestimmung des jeweiligen Peer Review im Wissenschaftssystem bestehen kann, und das in jedem Fall über den generellen Anspruch eines reliablen, validen und fairen Verfahrens hinaus geht.
- Zweitens: Welche expliziten und impliziten *Regeln* prägen das Verfahren? Damit sind nicht nur Regeln im Sinne von Bewertungskriterien gemeint, sondern dezidiert auch Regeln zur Personenkonstellation, zu Aufgaben-/Zuständigkeitsprofilen, zur Rollentrennung sowie zu Bewertungsordnungen und Begründungslogiken mitgedacht. Deskriptiv kann beispielweise unterschieden werden, ob es sich um formale Regeln, um rechtlich kodifizierte Regeln oder um informell eingebaute Konventionen handelt und wie diese begründet werden. Zudem ist herauszuarbeiten, in welchem Maße es sich bei all dem um Regeln handelt, die Standards für wissenschaftliche Qualität setzen - bspw. Originalität oder Robustheit? In diesem Kontext ist auch zu eruieren, ob spezifische Praxislogiken vorliegen - bspw. in Form von hochschul- und wissenschaftspolitischen Maßgaben, Zielen und Interessen, davon unabhängigen oder damit verknüpften reputationslogischen Ansinnen, Bestrebungen oder Konventionen bspw. zwecks Sichtbarkeit und Aufmerksamkeit, oder auch von ökonomischen Erwägungen, die mit Blick auf Mitteleinwerbung und Profit sowie einer günstigen Platzierung auf dem Forschungs- und Wissenschaftsmarkt relevant sein können.
- Drittens: In welchen *Infrastrukturierungen* manifestieren sich die mehr oder weniger regulierten Verfahren? In Anschluss an Susan Star und Karen Ruhleder sowie an Jörg Niewöhner lassen sich mittels der differenzierten Erfassung von Infrastrukturierungen im Sinne einer prozessualen Analyseperspektive weitere Einblicke in Bewertungspraktiken gewinnen; denn die Brücke zwischen Praktiken und ihrer verfahrensförmigen Organisation wird anhand von Infrastrukturierungen geschlagen. Zudem geben die Infrastrukturierungen Auskunft über den variablen Formalisierungsgrad der Verfahren, welcher u. a. mit Blick auf die spezifische Steuerungsfunktion relevant sein kann.
- Viertens: Was wird bewertet? Die Frage nach dem *Bewertungsgegenstand* ist nur vermeintlich einfach und taucht deshalb hier als letzte auf. Der Bewertungsgegenstand ist in viel geringerem Maße *ex ante* gegeben, als die Akteure im Feld dies wahrnehmen oder eingestehen. Es gilt deshalb danach zu fragen, *wie* der Gegenstand der Bewertung *durch das* und *in dem Verfahren* bestimmt wird; aber auch wie lange er unterbestimmt bleibt? Die Frage derart zu stellen, trägt der Tatsache Rechnung, dass die Dimensionen der Bewertung, welche den Gegenstand der Bewertung bestimmen, nicht deckungsgleich sein müssen, mit den Parametern, entlang derer die gültige Entscheidung getroffen wird. Für viele Verfahren ist es typisch, dass auf eine Einzelbewertung (bspw. von Anträgen auf Förderung) eine

vergleichende Bewertung und Entscheidung folgt, die die Bewertungsgegenstände erweitert (bspw. wenn Anträge in einer abschliessenden Gutachtersitzung unter Berücksichtigung regionaler Kriterien rangiert werden).

Auf diese Weise trägt die Frageheuristik dem Deweyschen Verständnis von Valuationsprozessen (Dewey, 1939) Rechnung, die ihren Ausgangspunkt an konkreten Problemen bzw. der Ungewissheit einer Bestimmung haben. Dementsprechend wird auch der spezifische Gegenstand der Bewertung erst im Zuge des Bewertungsverfahrens konstituiert. Gleichwohl ist diese Konzeption in der Empirie mit sehr heterogenen Fallbeispielen und dementsprechend diversen Selbstbeschreibungen von Gegenständen der Bewertung im jeweiligen Kontext konfrontiert. Heuristisch lassen sich grob folgende drei Varianten skizzieren:

- a. Etwas wird als Gegenstand der Bewertung bezeichnet, dass a priori bereits als eine physisch-materielle Einheit gegeben ist, die wiederum in einem Kontext bewertet wird, aber als Träger und Konvergenzpunkt des Bewertungsgeschehens unumgänglich ist (bspw. die Bewerber\*in im Berufungsverfahren).
- b. Etwas wird zu Beginn des Verfahrens als Gegenstand definiert bzw. eingegrenzt, z. B. im Zuge einer Selektion oder indem festgelegt wird, was als Gegenstand im Verfahren zugelassen wird (bspw. Artefakte, wie das formgerecht eingereichte Manuskript, der formal korrekt gestellte Antrag auf Forschungsförderung und die formgerechte Projektbeschreibung oder Bewerbung inklusive der personenbezogenen Informationen)<sup>3</sup>.
- c. Erst im Zuge des Verfahrens bildet sich sukzessive und jeweils einzelfallspezifisch heraus, was Gegenstand der Bewertung ist, so dass der Gegenstand erst im Verlauf des Bewertungsprozesses konturiert wird, beispielweise bei neuartigen Bewertungsverfahren, wie jenen im Rahmen der damals neuen Exzellenzinitiative oder aktuell der Verfahren zur Bewertung von Hochschulverbänden im Rahmen der Exzellenzstrategie (Möller, Antony, Hinze & Hornbostel, 2012; GWK, 2019).

Selbstverständlich sind in komplexen mehrstufigen Verfahren auch Mischformen denkbar, deren spezifische Ausgestaltung der Empirie zu entnehmen sein wird.

## 4 Verfahrensanalyse: Fragenspezifischer Vergleich

Diese Frageheuristik wird im Folgenden an konkreten Peer Review Prozessen zur Vergabe von Forschungsgeldern durchgespielt. Den Ausgangspunkt bildet das Fallbeispiel der Förderung von Sonderforschungsbereichen (SFB) durch die Deutsche Forschungsgemeinschaft (DFG). Dieser seit 1968 bestehende, langjährig erprobte zweistufige Peer-Review-Prozess beinhaltet verschiedenste Bewertungspraktiken (Expertenbegutachtung, Panel-Begutachtung, Vor-Ort-Begehung, Berichterstattung) und eignet sich daher als exemplarischer Fall, der umfassenden Einblick in die Komplexität der Peer-Review-Verfahren bietet. Eine bemerkenswerte Besonderheit dieses Peer-Review-Prozesses zur Entscheidung über die Einrichtung von Sonderforschungsbereichen besteht darin, dass diese als Verbundvorhaben konzipiert sind. Allerdings hat der Anteil der Drittmittel, die zur Finanzierung größerer Forschungsverbände mit längerer Laufzeit<sup>4</sup> vergeben werden im Zuge steigender Drittmittelquoten seit den 1990er Jah-

ren deutlich zugenommen (Klein, Kraatz & Hornbostel 2012, S. 164 f.; Deutsche Forschungsgemeinschaft, 2018, S. 188, 194), so dass die Analyse der Verbundförderung aussagekräftige Einblicke in die Entwicklung und den gegenwärtigen Stand von Förderformaten verspricht. In der vergleichenden Analyse werden außerdem zur gezielten Ergänzung weitere Fallbeispiele herangezogen. Die Auswahl der Förderprogramme zielt darauf ab, auch die im deutschsprachigen Raum etablierten Formate der Einzelprojekt- und der Personenförderung fragenspezifisch zu berücksichtigen. Derart bezweckt die komparative Überprüfung zu eruieren, inwiefern die Parameter der Analyseheuristik fallvergleichend zur Differenzierung und Typologisierung von Begutachtungsprozessen geeignet sind.<sup>5</sup>

## 5 Ziel und Zweck der Peer-Review-Verfahren

Um das Ziel zu bestimmen, gilt es, sich zu vergegenwärtigen, worüber mittels der Peer-Review-Verfahren entschieden wird. Selbstverständlich variieren je nach Förderlinie die Dauer, der finanzielle Umfang und die Adressaten der Förderung auf institutioneller sowie personeller Ebene: Beispielsweise wird beim Begutachtungsverfahren zur Einrichtung von Sonderforschungsbereichen über eine sehr langfristige (bis zu 12 Jahre) und im Gesamtvolumen vergleichsweise üppige und interdisziplinär angelegte Forschungsförderung befunden. Angesichts dessen ist das Peer-Review-Verfahren des SFB darauf ausgerichtet, fachlich kompetente und plausibel begründete Bewertungen hervorzubringen, welche diese langfristigen Förderentscheidungen hinreichend legitimieren.<sup>6</sup> Im Vergleich dazu wird bei der Einzelprojektförderung des Schweizer Nationalfonds (SNF) lediglich über eine drei- bis fünfjährige disziplinspezifische Förderung befunden, welche die Durchführung der geplanten Forschung ermöglichen soll.<sup>7</sup> Beim Emmy Noether-Programm der DFG wird über die mittelfristige Finanzierung einer Nachwuchsgruppe und ihrer Leitung (früher bis zu sieben Jahren, aktuell bis zu 6 Jahren) entschieden.<sup>8</sup> Das Emmy Noether-Programm bezweckt damit eine Weiterqualifizierung in Form des Erwerbs von Leitungserfahrung und der Möglichkeit zur fachlich eigenständigen Profilierung als Forscher\*in mit Blick auf das Karriereziel Professur.

Trotz unterschiedlicher Programme sticht der von allen geteilte Zweck ins Auge, fachwissenschaftlich kompetente und intersubjektiv abgesicherte Bewertungen vorzunehmen, die auch forschungs- und wissenschaftspolitisch legitimierbar sind, u. a. weil das Fachurteil eine realistische Abschätzung der jeweiligen Gelingensbedingungen, Zielerreichung und erfolgreichen Durchführung gewähren soll. Auffällig ist hier, dass die Verfahren immer auch dem Fortbestand und dem Nachweis der Funktionsfähigkeit der sie ausrichtenden Organisationen (DFG/SNF) selbst dienen; denn deren Aufgabe besteht darin, an der Schnittstelle zur Politik Forschungsmittel zu verteilen und zuverlässig Entscheidungen zu fällen. (Guston, 2000; Reinhart, 2012)

## 6 Regeln der Peer-Review-Verfahren

Mit Blick auf die zweite Frage, *welche expliziten und impliziten Regeln* die Peer-Review-Verfahren zur Vergabe von Drittmitteln prägen, lassen sich folgende Regeln entlang ihrer spezifischen Funktionen differenzieren:

- Regeln zur Legitimation des Urteils,
- Regeln zur Personenkonstellation,
- Bewertungsregeln im Sinne von Kriterien der Beurteilung,
- Verfahrensregeln und Regeln der Entscheidungsfindung (*decision making rules*).

Wiederum variiert je nach Förderprogramm, wie genau diese Regeln mit Blick auf die jeweilige Funktion ausgestaltet werden.

Bei der Bewertung der längerfristig einzurichtenden SFB werden zur *Legitimation der fachlichen Beurteilung* folgende Aspekte in Anschlag gebracht: Auffällig ist der hohe und auf die Bewilligung fokussierte Zeitaufwand von zwei vollen Arbeitstagen. Neben der Antragslektüre durch die Panelmitglieder wird hier in Form der Vor-Ort-Begehungen sowie der Klausursitzungen der Panelmitglieder und DFG-Berichtersteller\*innen beträchtlich in die systematische diskursive Aushandlung der Bewertung investiert. Dabei wird im Rahmen der Panelbegutachtung eine intersubjektive Nachvollziehbarkeit der individuellen Einschätzungen der Teilprojektgutachter\*innen angestrebt (Klein, Kraatz & Hornbostel, 2012, S. 165-166). Zudem steht am Ende der fachlichen Bewertung eine kollektiv geteilte Notenvergabe<sup>9</sup>. Derart zielt die Panelbegutachtung von SFBs auf eine konsensuale Bewertung ab. Die *Personenkonstellation* im SFB-Peer-Review-Verfahren ist durch eine starke Rollentrennung gekennzeichnet. Sie erfolgt zwischen vier Rollen: Erstens den Gutachter\*innen, welche die fachliche Expertise beisteuern sollen; zweitens den zwei Vertreter\*innen der DFG-Geschäftsstelle, die als Verfahrenshüter\*innen und Moderator\*innen fungieren; drittens den von der DFG berufenen Fachwissenschaftler\*innen, die als Berichtersteller\*innen das Verfahren begleiten und die Aushandlung der Bewertung überwachen; und viertens den zwei stimmberechtigten Politiker\*innen, die erst im letzten Verfahrensschritt – der Entscheidung im Bewilligungsausschuss – hinzugezogen werden. Die fachliche *Bewertung* erfolgt nach den *Kriterien* der Originalität der Forschungs idee und der Fragestellung. Darüber hinaus sind die Vorarbeiten und die Qualifikationen der Antragstellenden ein entscheidendes Kriterium (Olbrecht, 2014, S. 272 ff.). Wie Meike Olbrecht herausarbeitet, bieten diese Kriterien erheblichen Deutungsspielraum mit Blick auf ebenso schwer bestimmbare Maßstäbe wie Relevanz, Innovationspotentiale und zu erwartende Erkenntnisgewinne der Forschungsvorhaben. Dieser Befund trifft ebenso auf die Kriterien anderer Peer-Review-Verfahren zur Vergabe von Drittmitteln zu, bspw. auf die Bewertungskriterien beim Peer Review im Rahmen der Exzellenzinitiative (Möller, Antony, Hinze & Hornbostel, 2012; Hornbostel & Möller, 2015) oder des Emmy Noether-Programms (Böhmer, Hornbostel & Meuser, 2008)<sup>10</sup>. Bereits an der verwendeten Nomenklatur lässt sich ablesen, dass im Verfahren anhand der Bewertungskriterien auch Bewertungsmaßstäbe zu ihrer eigenen Bekräftigung inszeniert werden (Olbrecht, 2014, S. 272 ff.). Für eine vergleichende Verfahrensanalyse lassen sich die Terminologien der expliziten Bewertungskriterien somit selbst als Element der Regularien zur Legitimation des Qualitätsurteils lesen.

In den Regelungen zur Legitimation des Qualitätsurteils, der Personenkonstellation und den Kriterien, nach denen die Bewertung erfolgt, manifestieren sich folgende *Ver-*

*fahrensregeln* des SFB-Peer-Review-Prozesses: Es handelt sich um ein mehrstufiges Verfahren, das entsprechend der Rollentrennung arbeitsteilig organisiert ist. Die inhaltliche Bewertung der SFB-Anträge in der Vor-Ort-Begutachtung ist von der Entscheidung im Bewilligungsausschuss formal und auch personell weitgehend entkoppelt. Vermittels der im Verlauf der Verfahrensstufen und Sequenzen wechselnden Personenkonstellationen präformiert das Verfahren rollen- und aufgabenspezifische *Teilnahmestatus*. Die Gutachter\*innen sind an der Bewertung aktiv beteiligt und in ihrer fachlichen Urteilskompetenz voll akzeptiert. Von der abschliessenden Entscheidung sind sie jedoch ausgeschlossen. Die Mitarbeiter\*innen der DFG-Geschäftsstelle sind in der Bewertung zugegen, allerdings bleiben sie fachlich passiv, da sie nur begleitende, dokumentierende und teils moderative Funktionen erfüllen. Von der Entscheidung sind sie ebenfalls ausgeschlossen. Die Berichterstattenden sind als einzige sowohl in der Bewertung als auch in der Entscheidung zugegen. Sie übernehmen sowohl eine verfahrenshütende Funktion im Zuge der gemeinsamen Hervorbringung der Bewertung als auch eine synthetisch kommunizierende Funktion bei der Einigung auf eine konsensuale Bewertung im Panel; konkret in Form der kollektiv vertretenen Notengebung und deren Formulierung für das Entscheidungsgremium. Dort sind sie als Präsentierende aktiv und als berichterstattende Informant\*innen akzeptiert, allerdings nicht stimmberechtigt. Die Politiker\*innen sind von dem gesamten Bewertungsprozess ausgeschlossen und lediglich im Entscheidungsgremium zugegen. Auf Basis der ihnen vorliegenden Informationen sind sie dort aktiv stimmberechtigt und damit auch ohne fachliche Expertise gewissermaßen als funktionale Vertreter\*innen derjenigen, die Infrastrukturen und Gelder bereitstellen, an der Entscheidung beteiligt. Mit Blick auf den Teilnahmestatus kommt hinzu, dass die Verfahrensmäßigkeit durch wechselseitige Beobachtung der Verfahrensbeteiligten untereinander in Panel-Sitzungen, Diskussionen und Begehungen und mittels der Verfahrensbegleitung durch die Berichtersteller\*innen abgesichert wird. Die wechselseitige Kontrolle ist somit relativ stark formalisiert ausgeprägt. Bemerkenswert ist zudem, dass durch die direkte persönliche Begegnung von Gutachtenden, Verfahrensverantwortlichen und Antragstellenden die Anonymität der beteiligten Personen aufgehoben ist. Gleichwohl wird durch das kollektive Urteil in Form einer Note, die von der Gruppe als undifferenzierte Einheit vertreten wird und als Konsens kommuniziert wird, der Rückschluss auf Individualbewertungen einzelner Panel-Mitglieder verhindert.

Im Unterschied dazu erfolgt die *Legitimation des fachlichen Urteils* sowohl bei der Projektförderung des SNF als auch bei der Personalförderung der Nachwuchsgruppenleitung im Emmy Noether-Programm durch zwei bis drei jeweils unabhängig voneinander, schriftlich verfasste Gutachten von Fachkolleg\*innen, die von der Förderorganisation bestellt werden. Derart soll in Zusammenschau der Gutachten pro Antrag ein gegen einseitige Vorlieben oder Vorbehalte abgesichertes Fachurteil erzeugt werden. Beim Emmy Noether-Programm kommt ebenso wie bspw. bei der Förderung von Exzellenzclustern eine mündliche Präsentation der Antragstellenden hinzu. Diese sind somit, anders als beim SNF-Verfahren, nicht anonym, so dass auch die Informationen zum beruflichen Werdegang der Antragsteller\*innen explizit in die Begutachtung mit eingehen. Die *Personenkonstellation* dieser Verfahren umfasst neben den Antragsteller\*innen und den Fachgutachter\*innen auch die Mitglieder jener Förderorganisationen, bspw. Mitglieder des Hauptausschusses der DFG oder des Forschungsrates bzw. Referent\*innen des SNF. Somit ist die fachliche Begutachtung unterteilt in individuell

unabhängig voneinander verfasste Fachgutachten und eine gemeinsame Diskussion der Gutachten sowie der mündlichen Präsentation und des Diskussionsverlaufs im Entscheidungsgremium.

In diesen im Detail zwar unterschiedlichen Regeln zur Legitimation des Urteils, zur Personenkonstellation und zu den Bewertungskriterien manifestieren sich fallübergreifende *Verfahrensregeln und -merkmale*: Eine ausgeprägte Stufenförmigkeit, Rollendifferenzierung und Arbeitsteilung. Diese zeigt sich insbesondere darin, dass die beteiligten Akteure mitunter mit exklusiven Zuständigkeiten für bestimmte Dimensionen der Bewertung betraut sind: Beispielsweise kommt den Gutachter\*innen die Aufgabe zu, die fachliche Qualität und Relevanz einzuschätzen. Das Personal der Förderorganisationen übernimmt meist administrative und kontextualisierende Aufgaben. Den Mitgliedern der Entscheidungsgremien fällt auf Basis der vorangegangenen Bewertungsprozesse die relationierende Bewertung zu, in der meist expliziter als zuvor auch finanzielle und wissenschaftspolitische Aspekte mit abgewogen werden. Zudem sind in allen Fallbeispielen die fachwissenschaftliche Begutachtung und die Entscheidung über die Finanzierung somit formal ebenso wie personell entkoppelt und manifestieren sich für die jeweilige Verfahrensstufe und Sequenz als rollen- und aufgabenspezifisch differenzierte Teilnahmestatus. Auffällig ist auch die sequenzielle Organisation der verschiedenen Verfahrensschritte. Dabei wird im Zuge der Sammlung diverser fachnaher Einschätzungen im Sinne einer Perspektivenvielfalt, eine breite und komplexe Informationsbasis erzeugt, die dann im Zuge des Verfahrens, in der Relationierung zu anderen Bewertungsdimensionen sukzessive wieder reduziert und mit Blick auf die Entscheidung enggeführt wird.

## 7 Infrastrukturierungen der Peer Review Verfahren

Hinsichtlich der dritten Frage nach den *Infrastrukturierungen*, in denen sich die Regularien der Peer-Review-Verfahren niederschlagen, lassen sich folgende Differenzierungen vornehmen: Zunächst ist zwischen Schriftlichkeit und Mündlichkeit zu unterscheiden. Dabei sind Subunterscheidungen zu berücksichtigen: Einerseits zwischen Verschriftlichung in Form von Fließtext (bspw. bei Gutachten oder der Vergabe semantischer Noten) und numerischer Verschriftlichung in Form von Zahlen, Zeichen, Kennzahlen u. a. (bspw. numerische Noten oder Voten); andererseits zwischen mündlichen Diskussionen, Beratungen, Erörterungen in Begutachtungs-Panels und mündlichem Austausch mit Antragsteller\*innen in der unmittelbaren persönlichen Begegnung mit den an der Bewertung Beteiligten (Diskussionen, Frage-Antwort Dialoge etc.). Darüber hinaus kann der Formalisierungsrad auf einer Skala von informell zu formell, ggf. sogar bis hin zu rechtlich kodifizierten Regeln differenziert werden. Nicht zuletzt ist festzuhalten, welche prozeduralen Abläufe in physischer Kopräsenz oder zeitlich parallel und welche auch raum-zeitlich versetzt erfolgen und in welchem Maße die Prozesse analog bzw. digital erfolgen.

Alle erwähnten Peer-Review-Verfahren zur Vergabe von Drittmitteln basieren auf schriftlichen Anträgen zu Forschungsvorhaben, die inhaltlich teilstrukturiert sowie teilformalisiert sind. Sie alle enthalten neben freien Fließtextteilen auch skalierte Noten- oder Zeichensysteme, numerische und tabellarische Informationen zu Antragsteller\*innen sowie zur Finanzierung der geplanten Projekte. Die Verfahren zur Vergabe

von Drittmitteln sind zudem weitestgehend analog verankert. Auch wenn Dokumente und Informationen digital versendet werden, bedarf es für die aktenauglichen Dokumente meist zusätzlich einer postalischen Zustellung. Bisher erfolgt das Peer Review meist zumindest nicht ausschließlich über eine digitale Plattform oder Vergleichbares.<sup>11</sup> Unterschiede zeigen sich insbesondere mit Blick auf das Verhältnis und die Abfolge von Mündlichkeit und Schriftlichkeit im Begutachtungsverfahren. Im SFB-Peer-Review-Verfahren ist die Erarbeitung und Aushandlung der Bewertung im Zuge der Diskussionen im Panel und der Vor-Ort-Begehung von Mündlichkeit geprägt. Die Verschriftlichung erfolgt in diesem Verfahren konsequent im Nachgang der mündlichen Behandlung. Dabei nimmt der Formalisierungsgrad im Laufe des Verfahrens im Wechsel von mündlicher Diskussion anschließenden Verschriftlichungen in Beratungsvermerken, Protokollen, Notenvergabe und Entscheidungsvorlagen sukzessive zu. Diese Schriftdokumente fungieren somit als Instrument der sukzessiven Schließung eines mündlichen Gruppendiskussionsprozesses. Demgegenüber steht die weit verbreitete hybride Infrastrukturalisierung, die mit Blick auf Mündlichkeit und Schriftlichkeit gewissermaßen den umgekehrten Weg geht. Ausgehend von ausführlichen Fließtextverschriftlichungen hin zu einer mündlichen Diskussion, mündet sie in einer numerisch kondensierten Bewertung. Auch hier nimmt der Formalisierungsgrad im Zuge der Verknappung der schriftlichen Dokumente im Verfahrensverlauf zu.

## 8 Gegenstandsbestimmung der Peer Review Verfahren

Mit Blick auf die vierte Frage, *was im Zuge der beschriebenen Verfahrensweisen zum Gegenstand der Bewertung gemacht wird*, zeigt die Analyse, dass es sich keinesfalls um einen von Anfang bis Ende des Verfahrens fix zu identifizierenden Gegenstand handelt, sondern vielmehr um Gegenstandskonglomerate. Diese Konglomerate lassen sich in drei Komponenten untergliedern:

1. Die jeweiligen Textartefakte, d. h. Projektskizzen und Anträge, Gutachten, Sitzungsprotokolle, Vermerke, Berichte, Entscheidungsvorlagen;
2. Mündliche Einschätzungen, Diskussionsbeiträge und deren Relationierungen der Bewertenden und Entscheidenden (Gutachter\*innen in Panels, DFG Fachkollegiat\*innen, Berichterstatter\*innen);
3. Die schriftlich rhetorisch ausgearbeitete argumentative Überzeugungskraft der Antragsteller\*innen, die je nach Verfahren auch mündlich in Präsentationen und Diskussionen mit den an der Bewertung Beteiligten vorgeführt wird. Sie bezieht sich neben den inhaltlichen fachwissenschaftlichen Darlegungen insbesondere auch auf die strukturelle Passung zum Förderprogramm und ggf. auch auf die Darstellung der Persönlichkeit in der direkten Begegnung.

Mit Blick auf die Fallbeispiele ist deutlich, dass neben der qualitativen fachwissenschaftlichen Bewertung des Forschungsvorhabens anhand der Textartefakte und mündlichen Darlegungen sowohl bei der Verbundprojektförderung als auch bei der Personalförderung die geplante institutionelle Verortung des Forschungsvorhabens eine wichtige Rolle spielt. In diesen Förderlinien, beim Emmy Noether-Programm noch expliziter als beim SFB-Verfahren oder den Exzellenzclustern, werden zudem auch die Antragstellenden Personen (ob als Nachwuchsrgruppenleiter\*innen oder SFB- bzw. Cluster-Sprecher\*in-

nen) als Personen und in ihren bisherigen wissenschaftlichen Leistungen und Erfahrungen der Bewertung unterzogen. Hinzu kommen Form- und Stilaspekte der schriftlichen und mündlichen Darlegungen, die sich nicht nur in formaler Korrektheit, sondern darüber hinaus in situativer Angemessenheit, der Wortwahl, des Tonfalls, des Eingehens auf Fragen, der Kontextualisierung etc. niederschlagen. Auch diese können sich darauf auswirken, ob die an der Bewertung Beteiligten von der Relevanz und der Sinnhaftigkeit des beantragten Forschungsvorhabens überzeugt werden können.

Darüber hinaus werden im Zuge jener mehrstufigen Peer-Review-Verfahren zunehmend Meta-Bewertungen Teil des Gegenstands, beispielsweise wenn in den Diskussionen zu den Gutachten Stellung genommen wird und diese mit Blick auf ihre Trefflichkeit und Tauglichkeit als Bewertungen bewertet werden oder aber wenn Empfehlungen von Referent\*innen, Panelmitgliedern, Vermerke von Geschäftsstellenmitarbeiter\*innen oder Entscheidungsgrundlagen ihrerseits bewertet werden. Diese Aneinanderreihung von Bewertungen und Meta-Bewertungen, geht damit einher, dass zwar nicht unbedingt in physischer Kopräsenz (wie im Panel), sondern auch raum-zeitlich versetzt (Gutachten, Vermerke, Empfehlungen) die Bewertenden selber wiederum in ihrer bewertenden Praxis einer Bewertung unterzogen werden. In der Zuspitzung läuft diese Anhäufung von Metabewertungen auf das hinaus, was Hirschauer als „Beurteilungspanoptismus“ bezeichnet (Hirschauer, 2019, S. 286 ff.).

Dabei fällt auf, dass in Peer-Review-Verfahren zur Drittmittelvergabe der Gegenstand der Entscheidung meist nicht identisch mit dem Gegenstand der Bewertung ist, weil die Begutachtungsprozesse häufig von den Entscheidungsprozeduren entkoppelt sind. Nicht zuletzt bedingt die sequenzielle Organisation der Verfahren, dass sich die Dokumentensammlung, über die bewertet und entschieden wird, im Laufe des Peer-Review-Verfahrens wandelt. Die ausgeprägte Rollendifferenzierung und Arbeitsteilung der Verfahrensbeteiligten bedingt überdies, dass in den von der inhaltlichen Fachbegutachtung separierten Entscheidungsgremien, auf der Grundlage von Entscheidungsvorlagen mitunter auch wissenschaftspolitische und mit Blick auf den gesamten Zuständigkeitsbereich der Förderorganisation verteilungsstrategische Gesichtspunkte eine Rolle spielen.

## 9 Erweiterte Prüfung der Heuristik

Angesichts des Befunds, dass sich bei den erörterten Fallbeispielen aus der Forschungsförderung die Gegenstände der Bewertung erst im Zuge des Verfahrens konstituieren und es sich somit um im Verfahren herausgebildete Gegenstandskonglomerate handelt, wird im Folgenden ein kursorischer Blick auf die Verfahren des Zeitschriften-Peer-Review und des Peer Reviews im Rahmen von Berufungsverfahren gerichtet. Dies dient dazu, die Beständigkeit und Tauglichkeit der entwickelten Heuristik zu eruieren und das Potential der fallvergleichenden Analyse für Peer-Review-Prozesse jenseits der Drittmittelvergabe auszuloten. Aufgrund des begrenzten Platzes begnügen wir uns damit, diese Prüfung lediglich anhand der Eckpunkte der Heuristik, d. h. für die *Ziele und Zwecke* der Verfahren sowie für die *Gegenstandsbestimmung* im Verfahren vorzunehmen.

Ziele des Zeitschriften-Peer-Reviews sind die Kalibrierung der Lesezeit einer Disziplin sowie allgemein die wissenschaftliche Qualitätssicherung der publizierten Fachartikel (Hirschauer, 2005, S. 53 ff.; 2019, S. 227). Ziel eines Berufungsverfahrens ist es, die bestmögliche Passung von Entwicklungszielen der Hochschule und Fakultät mit der

langfristig angestellten Person zu erzielen, wie die Studien von Bernd Kleimann und Maren Klawitter zeigen (Klawitter, 2015; Kleimann, In der Smitten & Klawitter, 2015; Kleimann & Klawitter, 2017a; 2017b). Die Berufung dient damit dem Zweck der Profilbildung der Hochschule sowie der Aufgabenerfüllung ihres Personals in Forschung und Lehre.

Mit Blick auf den Gegenstand wird im *Zeitschriften-Peer-Review* anhand der Gutachten die inhaltliche und formale Qualität des Manuskripts, seine Passung zur Zeitschrift, zur Ausgabe, die Angemessenheit und Gültigkeit des im Manuskript erhobenen Geltungsanspruches und der im Gutachten vorgebrachten Beurteilung und Kritik bewertet. Somit handelt es sich auch hier um ein Gegenstandskonglomerat, das der Bewertung unterzogen wird. Diese richtet sich auf das eingereichte Manuskript und den darin erhobenen Geltungsanspruch der Autor\*innen, etwas im Rahmen dieser Zeitschrift und ihrer Leserschaft Passfähiges und Publikationswürdiges mitzuteilen. Im *Berufungsverfahren* werden die Werdegänge und Merkmale der Bewerber\*innen auf die ausgeschriebene Professur, d. h. die Eigenschaften bewertet, an denen ihre Eignung, Befähigung und ihr Leistungspotential abgeschätzt werden. Somit ist hier die Person in Hinblick auf ihre Passung zu den hochschul- und ggf. fakultäts- und institutsspezifischen Entwicklungen Gegenstand der Bewertung. Mit Blick auf die im Verfahren zusammengetragenen und eingeschätzten Informationen, entpuppt sich allerdings auch die Person als ein Konglomerat aus nachgewiesenen, beschriebenen, dargestellten und zugeschriebenen Merkmalen. Sie wird als ein Bündel aus a) manifesten, d. h. in Publikationen, eingeworbenen Mitteln, dokumentierten Forschungsleistungen etc.; b) einem biographischen Verlauf, der Aufenthalts- und Arbeitsorte listet, die sich mit Blick auf qualitätsdifferenzierende Zuschreibungen wie Reputation der Einrichtung, Internationalität, Wechselhäufigkeit etc. lesen und deuten lassen und c) der Persönlichkeit der Bewerber\*innen, die grob als soziales Verhaltensprofil mit Blick auf kommunikative, kooperative, argumentative, lehrdidaktische Fähigkeiten und eigene Interessen und nicht zuletzt der Gewinnungswahrscheinlichkeit der Bewerber\*in gefasst und bewertet. In vergleichender Perspektive hat dabei das Zeitschriften-Peer-Review mit manchen Verfahren zur Drittmittelvergabe, bspw. mit der Projektförderung des SNF die Anonymisierung der Einreichenden gemein. Hingegen teilen die Berufungsverfahren mit manchen Verfahren zur Vergabe von Fördergeldern die persönliche Präsentation der Bewerber\*innen.

## 10 Zwischenbilanz – Elemente des Peer Review

Trotz der Unterschiedlichkeit der dargelegten Fallbeispiele werden im Vergleich somit Überschneidungspunkte und Gemeinsamkeiten sichtbar: Zum einen der formale Befund, dass trotz divergierender Ziele, die Peer Review Verfahren immer selbst ein konstitutiver Bestandteil für die Bestimmung der Gegenstandskonglomerate sind, die hier einer Bewertung unterzogen werden. Zum anderen kristallisieren sich im Zuge der vergleichenden Verfahrensanalyse der Bewertungspraktiken spezifische Tätigkeiten heraus, die sich als elementare Bestandteile von Peer-Review-Verfahren identifizieren lassen:

### *Postulierende Tätigkeit*

Am Beginn des Prozesses steht ein Postulat. Die Einreichung eines Manuskripts, eines Förderantrags sowie der Bewerbung verstehen wir als postulierende Tätigkeit im Sinne

eines Antrags auf Genehmigung der damit verbundenen Ansprüche. Genauer gesagt betrachten wir den Anspruch auf Publikationsfähigkeit, den Anspruch auf Förderfähigkeit für ein Projekt oder eine Person und den Anspruch, dem ausgeschriebenen Stellenprofil zu entsprechen und es kompetent ausfüllen zu können, als Forderung nach Akzeptanz des Gesuchs. Diese Postulate schaffen den Anlass für die Bewertung, ihre Durchführung und ihre Organisation in Verfahren, um später eine evaluationsbasierte Entscheidung zu ermöglichen.

#### *Konsultative Tätigkeit*

Zudem sind konsultative Aktivitäten wesentlich für Peer Review. Die Prozesse, in denen das Manuskript, der Vorschlag oder die Bewerbung und gegebenenfalls die persönliche mündliche Präsentation überprüft werden, verstehen wir als Beratungspraxis. Die schriftlichen Überprüfungen, Notizen und mündlich geäußerten Schätzungen zu den Postulaten werden als konsultative Gutachten eingeholt, um eine kompetente Bewertung zu gewährleisten.

#### *Dezisionale Tätigkeit*

Ebenso konstitutiv für den Peer-Review-Prozess ist, dass er eine Entscheidung zustande bringt. Der Verfahrensschritt, in dem die berechtigten Teilnehmer\*innen des jeweiligen Verfahrens eine verbindliche Entscheidung über die Annahme oder Ablehnung des Postulats treffen – sei es auf der Grundlage von Noten, Rankings, Volltextmitteilungen wie Berichten oder mündlichen Präsentationen – bezeichnen wir als entscheidungstreffende Tätigkeit.

#### *Administrative Tätigkeit*

Unerlässlich sind zudem verwaltende Aktivitäten. Die Prozesse, in denen Postulate empfangen und ihre Behandlung eingeleitet und koordiniert werden, begreifen wir als administrative Tätigkeit. Diese Aktivitäten übernehmen in der Regel jene Organisationen, die Peer-Review-Prozesse einsetzen. Dementsprechend werden sie von deren Vertreter\*innen wie Herausgeber\*innen, editorialen Manager\*innen oder den Mitarbeiter\*innen der Fördereinrichtungen, Verlagen oder universitären Personalabteilungen durchgeführt. Die Verwaltungstätigkeit als eigenständige Tätigkeit zu definieren, trägt der Tatsache Rechnung, dass Peer-Review-Verfahren gewöhnlich in einem bestehenden organisatorischen Rahmen erfolgen.

#### *Diskutierende Tätigkeiten*

Hinzu kommen diskursive und erörternde Aktivitäten. Wir definieren die Prozesse, in denen die Gutachten, Notizen, Einschätzungen, die zu den Bewertungen beitragen, ausgetauscht, kommentiert und diskutiert werden als Diskusstätigkeiten. Diese Diskussionspraktiken finden z. B. in Redaktionen, Herausgeber\*innengremien, Gutachter\*innensitzungen und Panels bzw. Entscheidungsgremien oder Berufungskommissionen statt und können schriftliche Gutachten im Fließtext sowie Memos, Sprechnotizen oder numerische bzw. zeichenbasierte Bewertungen und mündliche Beiträge beinhalten.

#### *Präsentierende Tätigkeiten*

Ergänzend sind darbietende und vorzeigende Aktivitäten zu differenzieren. Die Prozesse, in denen die Postulierenden ihren Projektvorschlag, ihr Profil und ihre Anfrage persönlich vorstellen und erläutern, einschließlich der Beantwortung der damit verbundenen Fragen, fassen wir als präsentierende Tätigkeiten. Bei diesen Aktivitäten sind die

Postulierenden in den Bewertungsprozess mit eingebunden. Diese Aktivitäten sind in meist mehrstufigen Verfahren für Bewerbungen und personalisierte Zuschüsse sowie für Projekt- und Verbundprojektfinanzierungen zu finden, z. B. wenn im Rahmen der Begutachtung oder Evaluation Vor-Ort-Begehungen oder Besuche erfolgen. Meistens sind diese Tätigkeiten im Nachgang einer mehr oder weniger intensiven Einschätzung oder Bewertung schriftlicher Postulate veranschlagt.

### *Beobachtende Tätigkeiten*

Zu unterscheiden sind überdies observative Aktivitäten, die auch der Verfahrenskontrolle dienen. Die Prozesse, in denen das Peer-Review-Verfahren mit Blick auf die Einhaltung der jeweiligen Aufgaben, Rechte und Pflichten jeder Rolle, des Verfahrensablaufs und des Gesamtziels der Prozedur begleitet und gehütet wird, werden von uns als beobachtende Aktivitäten eingestuft. Diese Aktivitäten sind oft mit der Verpflichtung verbunden für die Entscheidungsgremien zu dokumentieren oder Bericht zu erstatten.

### *Moderierende Tätigkeiten*

Darüber hinaus bedingt die Ausdifferenzierung der Tätigkeiten einen Bedarf an Moderation, welche die situative Rahmung gewährleistet, in dem die Einzelaktivitäten zueinander in Beziehung gesetzt werden. Das geregelte Verfahren als solches erfordert moderierende Tätigkeiten, welche die Prozesse begleiten und bei Bedarf erläutern und erklären. Je komplexer und mehrstufiger sie sind, desto stärker muss das Zusammenspiel der verschiedenen Aktivitäten im laufenden Verfahren explizit abgestimmt und koordiniert werden, z. B. indem die Beteiligten in das Verfahren eingeführt und durch dieses begleitet werden. Dies kann auch die Übernahme von Sitzungsleitungen bei Diskussionen oder Präsentationen von Antragsteller\*innen oder Bewerber\*innen beinhalten.

Auf Grundlage der identifizierten elementaren Bestandteile von Peer-Review-Verfahren lässt sich wiederum der Blick für eine typologisierende Differenzierung der Verfahren öffnen. Angesichts der acht unterschiedenen Elemente des Peer Review lassen sich nun einerseits jene davon identifizieren, die als notwendige Bestandteile ein Minimalverfahren kennzeichnen: Postulierende, konsultative, dezisive und administrative Tätigkeiten. Andererseits kommen weitere Elemente modulhaft hinzu, die eine Anpassung an weitere Verfahrensziele und Kontexte leisten können: Diskutierende, präsentierende, observative und moderierende Tätigkeiten konturieren alle zusammengenommen ein maximales Verfahren (siehe *Tabelle 1*).

*Tabelle 1: Verfahrenskonstellationen des Peer Review*

<b>Minimalverfahren</b>	<b>Zusatzmodule in Richtung Maximalverfahren</b>
Postulat	Diskussion
Administration	Präsentation
Konsultation	Observation
Dezision	Moderation

*Quelle:* Eigene Darstellung.

Derart können die Tätigkeiten auch als verschiedene Elemente gesehen werden, die ähnlich einem Baukastensystem je nach Ziel und Zweck des Verfahrens, Steuerungs- und Legitimationsbedarf unterschiedlich kombiniert werden können. Dabei sind diverse Anteilsverhältnisse vorzufinden, die bedingen, welche Tätigkeiten das Verfahren in

welchem Maße prägen oder gar dominieren. Beispielsweise fällt das Fachurteil von Expert\*innen in den Rahmen der konsultativen Tätigkeiten, so dass bei minimalistischen Peer-Review-Verfahren das Expert\*innenurteil einen größeren Stellenwert im Evaluierungsarrangement einnimmt. Dessen relatives Gewicht nimmt bei aufwändigeren und dadurch stärker modularisierten Verfahren tendenziell ab, da dort auch moderierende, präsentierende, beobachtende und diskutierende Aktivitäten im Verfahren integriert sind. Diese Binnendifferenzierung ist häufig zu beobachten, wenn zusätzlich zur fachlichen Qualitätssicherung die Scharnierfunktion zur Vermittlung unterschiedlicher Steuerungsinteressen bspw. von Wissenschaftsorganisationen und politischen Akteuren gewichtig ist. Zudem steigt mit zunehmender Komplexität der Verfahren auch der Bedarf, die unterschiedlichen Aktivitätsmodule miteinander zu artikulieren, was eine stärkere Formalisierung der Maximalverfahren u. a. in der Prozessgestaltung bedingt. Diese geht dann bevorzugt auch mit zunehmend formaler Infrastrukturalisierung mitunter gar mit Standardisierung einher.

## 11 Fazit: Evaluationspolitik der Wissenschaft - Regierung durch Peer Review

Die entwickelte Heuristik zielte darauf ab, Peer-Review-Formate als Bewertungsprozesse zu erschließen, um sie einer vergleichenden Verfahrensanalyse zugänglich zu machen. Die Heuristik fokussiert dementsprechend verfahrensrelevante Aspekte und ermöglicht dadurch viererlei: Erstens können Peer-Review-Verfahren nach ihrem *spezifischen Zweck* differenziert werden; zweitens ist der *Bewertungsgegenstand* nicht zwingend von Beginn an als scheinbar gegebener gesetzt; drittens erlaubt der weit gefasste Begriff von *Regeln*, die verfahrensförmige *Infrastrukturierung* der Praxis zu analysieren; und viertens ermöglicht dies, die häufig im Laufe des Verfahrens wechselnden Teilnahmestadien, bspw. anhand der Personenkonstellationen, Arbeitsteilung und Rollentrennung zu beschreiben und zu erfassen, ohne sie dabei statisch Einzelakteuren dauerhaft zuschlagen zu müssen. Derart werden der Prozessualität und Sequenzialität jener Verfahren konsequent Rechnung getragen.

Mittels der Analyseheuristik können somit Gemeinsamkeiten und Unterschiede der vielfältigen Peer-Review-Verfahren identifiziert werden, so dass greifbar wird, wie die Vielfalt, innerhalb der allen gemeinsamen Konzeption als Peer Review möglich ist. Erkennbar wird dabei auch, dass eine solche Diversität in der Einheit gewissermaßen Bedingung der Möglichkeit jener auch historisch zu konstatierenden „Metastabilität des Peer Review“ sein muss. Es scheint nachgerade das kennzeichnende Merkmal wissenschaftlicher Evaluierungsarrangements zu sein, dass sie auf den ersten Blick den Anschein erwecken müssen, einheitlich und wissenschaftsspezifisch zu sein, damit sie als Peer Review gelten können. Die Diversität innerhalb dieser Einheit der Peer-Review-Form offenbart hingegen, dass von Einheitlichkeit allenfalls in Bezug auf einzelne modulare Tätigkeiten gesprochen werden kann.

Angesichts der diversen Aktivitätsmodule und deren vielfältiger Fremddifferenzen ist zudem evident, dass es im Peer Review keinesfalls nur darum geht, zu bewerten. Die Kalibrierung der Lesezeit einer Disziplin, die fachkompetente intersubjektiv abgesicherte Qualitätsbewertung, die in Themensetzung, personeller Ausstattung und finanzieller Investition wissenschaftspolitisch legitimierbar ist, die Personalpassung mit

Blick auf fachliche und strukturelle Faktoren im Verhältnis zu den Zielen der Hochschulentwicklung und weitergedacht für Promotions- oder Habilitationsverfahren die Vergabe qualifikationsverbürgender Titel; all das sind exemplarische Leistungen, die dem Peer Review überantwortet werden. Neben der evidenten Steuerungsfunktion ist damit die Legitimationsfunktion augenfällig zentral und zwar nicht nur nach innen gegenüber der wissenschaftlichen Gemeinschaft, sondern auch nach außen, nicht zuletzt gegenüber der Politik. In diesem Spannungsfeld ist eine einheitliche Adressierbarkeit nötig, da es um die Verwendung öffentlicher Gelder für Forschung geht und zudem die Zugänglichkeit von daraus resultierendem als gesichert geltendem Wissen zur Debatte steht. Insofern liegt von vielfältigen Seiten – Wissenschaft, Politik, im Fall von Wissenschaftsverlagen auch Wirtschaft – ein Steuerungsinteresse vor, das in einer geteilten Terminologie ausgehandelt werden muss, um es handhabbar im Sinne von regierbar zu machen. Das vereinheitlichende *Label* „Peer Review“ gewährleistet für alle Akteure eine Erkennbarkeit, die nicht nur Steuerungsinteressen vermittelt, sondern in erheblichem Maße Legitimation für eine Regierung durch Peer Review verbürgt.

Dieses zeigt sich einerseits an der Geschichte des Begriffs „Peer Review“. Als verbreitete Bezeichnung für wissenschaftliche Evaluierungsarrangements taucht dieser Terminus nämlich erst in den 1970er Jahren auf und vermittelt erstmals zwischen neuen politischen Steuerungsinteressen und wissenschaftlichen Autonomieforderungen (resp. Selbststeuerungsinteressen) (Baldwin, 2018). Hier findet sich der historische Moment, in dem vielfältige wissenschaftliche Evaluierungsarrangements durch eine vereinheitlichende Terminologie legitimatorisch gewendet werden. Darüber hinaus sind die in der Folge dessen komplexer werdenden Verfahren eine Reaktion darauf, dass das Peer Review selbst wegen seiner Steuerungsfunktion erhöhtem Legitimationsdruck ausgesetzt ist. Reflexive Modernisierung und Wissensgesellschaft sind die Schlagworte, mit denen der gesellschaftliche Kontext seit den 1970er Jahren bezeichnet werden kann, der diesen Druck erzeugt (Weingart, 2001). Derart ist Peer Review nicht nur selbst Gegenstand von Politik und insbesondere von Evaluationspolitiken geworden, sondern hat dadurch auch eine immense Verbreitung seiner Anwendung und Nutzungsformen sowie Ausdehnung seiner Wirkmächtigkeit erfahren. Folgerichtig eröffnen sich im Ausblick für die Wissenschaftsforschung Möglichkeiten, den Themenkomplex um Qualitätssicherung, Ressourcenzuweisung und Expertise aus der leitenden Unterscheidung von Selbst- und Fremdsteuerung herauszulösen.

Angesichts des zunehmenden Legitimationsdrucks ließe sich die Frage weiterverfolgen, inwiefern eine zunehmende Ausdifferenzierung der Verfahren durch zusätzliche Aktivitäten deren Komplexität übermäßig steigert. Einerseits erfordern die Verfahren dadurch immer mehr Vorkenntnisse und Einarbeitung, bis hin zu Professionalisierung, die bspw. am Element der konsultativen Tätigkeiten anhand der Notwendigkeit von Gutachter\*innen-Trainings diskutiert werden. Extremes Beispiel hierfür ist die Einführung von sog. „impact case studies“ im Research Excellence Framework Grossbritanniens, wo sowohl die Evaluierten als auch die Evaluierenden in vorgängigen Testläufen regelrecht auf das Verfahren hin „kalibriert“ wurden (Manville et al., 2015). Andererseits droht die zunehmende Verfahrenskomplexität auch die Einheit dessen zu sprengen, was unter Peer Review gefasst wird und was seine Meta-Stabilität bedingt. Alternative Verfahren zum Peer Review erscheinen dadurch zumindest auf den ersten Blick effektiver und effizienter, so wie das gegenwärtig bspw. in der Diskussion um Losverfahren in der Forschungsförderung sichtbar wird. Zudem tritt auch das spezifisch wissenschaftliche am Peer Re-

view dann in den Hintergrund zugunsten der vielfältigen Fremddifferenzen, die Peer-Review-Formate ermöglichen und die insbesondere die Scharnierfunktion zum Politischen im Sinne eines Adaptationsvermögens in den Vordergrund treten lassen.

Als Schlussfolgerung unserer fallvergleichenden Analyse ist jedoch die Identifikation konstitutiver Aktivitäten und deren zielgerichtete modularisierte Zusammenstellung u. a. im Zuge von Maximalverfahren von vorrangiger Bedeutung: Die praxeologische Analyse befördert aus dem empirischen Material zu Tage, in welchem Maße Peer Review nicht nur in Evaluationspolitiken eingebettet ist, sondern auch mit politischen Gestaltungsfragen verwoben ist. In dieser Perspektive lassen sich die vielfältigen Evaluierungsarrangements des Peer Review mit Pierre Rosanvallon (2018) auch als verschiedene Regierungsvarianten in der Wissenschaft begreifen. Diese unterscheiden sich danach, wie in den divergierenden Verfahren die verschiedenen Steuerungsinteressen prozessualisiert und in sequenzialisierte und formalisierte Abläufe überführt werden, welche Tätigkeiten dabei wem zufallen, kurzum welche Verfahrensförmigkeit welche Teilnahmestatus, d. h. Repräsentations- und Beteiligungsformen zur Entscheidung über die Verteilung knapper Ressourcen vorsieht. Derart lassen sich die Evaluierungsarrangements des Peer Review als Regierungsarrangements der Wissenschaft begreifen. Im Zuge dessen werden wissenschaftliche Anforderungen an die Qualitätsprüfung innerhalb des Peer Review bspw. anhand des (langjährig erprobten) Verfahrenselements der Konsultation von Fachleuten an vorrangig politisch geprägte Aushandlungs-, Bewertungs- und Entscheidungsprozeduren wie bspw. die diskursiven und observativen Verfahrenselemente anschlussfähig gemacht. Die Kombination von auf das Spezialist\*innenurteil verweisenden und auf Einbindung, Aushandlung sowie auf die Kontrolle und Gewährleistung einer akuraten Rollentrennung und kompetenzgeleiteten Arbeitsteilung ausgerichteten Verfahrensmo­du­len steht sinnbildlich dafür, wie Peer Review als ein Adapter fungiert, der wissenschaftliche und politische Qualitäts- und Legitimationserfordernisse ineinander überführt und füreinander kompatibel macht. Eine gouvernementale Nomenklatur (Foucault, 2006) erlaubt hier das verwobene Verhältnis von Regierenden und Regierten differenzierter bspw. als Momente von Repräsentation, Partizipation, Rollentrennung, aber auch von Veridiktion, Lesbarkeit und Responsivität zu untersuchen. Nicht zuletzt würden dadurch wissenschaftliche Evaluierungsarrangements noch deutlicher als Evaluationspolitik erkennbar.

## Anmerkungen

- 1 Siehe dazu die Rezeption der Literatur in Weller (2001), Hirschauer (2005), Reinhart (2012), Neidhardt (2016).
- 2 Darüber hinaus diagnostiziert Hirschauer dem Zeitschriften-Peer-Review Strategien der Kommunikationsintensivierung, die auf die Steigerung der Aufmerksamkeit, der Sachlichkeit, der Meinungsverschiedenheit und der Disziplinierung der Autorschaft zielen (Hirschauer, 2019, S. 294).
- 3 Mit Blick auf Fallanalysen ließe sich auch hier wieder die Frage produktiv machen, in welchem Maße diese Festlegung als eine Wertzuschreibung erfolgt und/oder in welchem Maße sie im Modus der Wertabwägung getroffen wird. Aufgrund des erläuterten mangelnden Generalisierungspotentials wird diese Frage allerdings hier zurückgestellt.
- 4 Darunter fallen in Deutschland sowohl die von der DFG im Rahmen sogenannter „koordinierter Programme“ – neben Sonderforschungsbereichen auch Forschungsgruppen, Forschungszentren, Schwerpunktprogramme und Graduiertenkollegs – als auch die seit Anfang der 2000er Jahre im Zuge der Exzellenzinitiative bzw. seit 2019 der Exzellenzstrategie von der öffentlichen Hand verausgabten Forschungsgelder.

- 5 Die Informationen zu den Fallbeispielen speisen sich aus verschiedenen Einzelfallstudien und der dazugehörigen Literatur. Insbesondere bei den Drittmittelverfahren wird auf umfangreiche empirische Studien im Umfeld des Instituts für Forschungsinformation und Qualitätssicherung (iFQ), heute Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), zu Sonderforschungsbereichen, zum Emmy Noether-Programm und zur Projektförderung des Schweizerischen Nationalfonds (SNF) zurückgegriffen, die ergänzt durch Webseiten- und Dokumentenanalyse sekundäranalytisch ausgewertet werden.
- 6 Auffällig ist, dass auch strukturelle Faktoren und wissenschaftspolitische Erwägungen eine gewichtige Rolle spielen. Aufgrund der langfristig angelegten Förderung (4+4+4) wird die Förderungswürdigkeit auch mit Blick auf inner- und interdisziplinäre Entwicklungen, die Relevanz im internationalen Vergleich und die Komplementarität zu bereits bestehenden Schwerpunkten eingeschätzt und abgewogen (Olbrecht, 2014).
- 7 Das Verfahren des Schweizerischen Nationalfonds zielt auf eine möglichst gut abgesicherte und somit fachwissenschaftlich und forschungspolitisch legitimierbare, da angemessene und sachgerechte Entscheidung über die Vergabe von finanziellen Mitteln zur Projektförderung (Reinhart, 2012).
- 8 Das Begutachtungsverfahren des Emmy Noether-Programms ist als Personenförderung für promovierte Wissenschaftler\*innen angelegt. Die Antragsstellenden beantragen die Einrichtung und Leitung einer Nachwuchsgruppe zu einer spezifischen Problemstellung, die an eine etablierte Forschungseinrichtung angegliedert ist (Deutsche Forschungsgemeinschaft, 2019)
- 9 Zu den Effekten gruppendynamischer Prozesse in der Panelbegutachtung bspw. auf die kollektive Urteilsfindung in konkurrenzialen Verfahren und die Notenvergabe siehe auch Klein, Kraatz & Hornbostel, 2012, S. 170-181.
- 10 Die fachliche Bewertung des Forschungsvorhabens für eine Nachwuchsgruppe wird anhand von Kriterien der Originalität und Konzeption des Forschungsprojektes eingeschätzt. Darüber hinaus wird auch die Passung des Forschungsvorhabens und der/des sich bewerbenden Nachwuchsleiter\*in/s zum Programm bewertet, die u. a. trotz der mit Blick auf das wissenschaftliche Alter relativ jungen Zielgruppe bereits Anforderungen wie internationale Forschungserfahrung und hochrangige Publikationen miteinbezieht. Die Bedeutung dieser Kriterien zeigt sich insbesondere in dem nach Meinung der Antragssteller\*innen sehr bedeutsamen Merkmal der persönlichen Publikationsleistung (inklusive der Reputation der Zeitschriften) und bisheriger Forschungsaktivitäten (Böhmer, Hornbostel & Meuser, 2008, S. 49).
- 11 Sowohl bei den genannten Verfahren als auch generell ist allerdings eine Verschiebung der Kommunikation auf digitale Plattformen deutlich erkennbar.

## Literatur

- Baldwin, Melinda (2018). Scientific Autonomy, Public Accountability, and the Rise of “Peer Review” in the Cold War United States. *Isis*, 109 (3), 538-558. <https://doi.org/10.1086/700070>.
- Biagioli, Mario (2002). From book censorship to academic peer review. *Emergences: Journal for the Study of Media & Composite Cultures*, 12 (1), 11-45. <https://doi.org/10.1080/1045722022000003435>.
- Bogusz, Tanja (2018). *Experimentalismus und Soziologie: Von der Krisen- zur Erfahrungswissenschaft*. Frankfurt: Campus Verlag.
- Böhmer, Susan, Hornbostel, Stefan, & Meuser, Meuser (2008). Postdocs in Deutschland: Evaluation des Emmy Noether-Programms. *iFQ-Working Paper No.3*. Bonn.
- Deutsche Forschungsgemeinschaft (2018). *Jahresbericht 2018*. Aufgaben und Ergebnisse. Bonn. Verfügbar unter: [https://www.dfg.de/download/pdf/dfg\\_im\\_profil/geschaeftsstelle/publikationen/dfg\\_jb2018.pdf](https://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/dfg_jb2018.pdf) / [17. März 2020].
- Deutsche Forschungsgemeinschaft (2019). *Emmy Noether-Programm*. Bonn. Verfügbar unter: [https://www.dfg.de/foerderung/programme/einzelfoerderung/emmy\\_noether/](https://www.dfg.de/foerderung/programme/einzelfoerderung/emmy_noether/) [29. November 2019].
- Dewey, John (1939). *Theory of Valuation*. Chicago: University of Chicago Press.
- Foucault, Michel (2006). *Sicherheit, Territorium, Bevölkerung*. (C. Brede-Konersmann & J. Schröder, Übers.). Frankfurt am Main: Suhrkamp.
- Goffman, Erving (1980). *Rahmen-Analyse: Ein Versuch über die Organisation von Alltagserfahrungen* (10. Aufl.; H. Vetter, Übers.). Frankfurt am Main: Suhrkamp Verlag.

- Goffman, Erving (1986). *Interaktionsrituale: Über Verhalten in direkter Kommunikation* (1. Aufl.; R. Bergsträsser & S. Bosse, Übers.). Frankfurt am Main: Suhrkamp Verlag.
- Goffman, Erving (2005). *Rede-Weisen: Formen der Kommunikation in sozialen Situationen* (H. Knoblauch, C. Leuenberger, & B. Schnettler, Hrsg.). Konstanz: UVK Verlagsgesellschaft.
- Guston, David H. (2000). *Between Politics and Science: Assuring the Integrity and Productivity of Reseach*. (2 Aufl.) Cambridge; New York: Cambridge University Press.
- GWK-Bonn (2019). *Die Exzellenzstrategie und ihre Vorgängerin, die Exzellenzinitiative*. Bonn. Verfügbar unter: <https://www.gwk-bonn.de/themen/foerderung-von-hochschulen/exzellenzstrategie-exzellenzinitiative/> [29. November 2019].
- Heintz, Bettina (2018). Von der Allmacht der Zahlen und der Allgegenwart des Bewertens. *Soziologische Revue*, 41 (4), 629-642.
- Hirschauer, Stefan (2005). Publierte Fachurteile. Lektüre und Bewertungspraxis im Peer Review. *Soziale Systeme*, 11(1), 52-82.
- Hirschauer, Stefan (2015). Oral Communication in Journal Peer Review. *Human Studies* 38: 37-55.
- Hirschauer, Stefan (2019). Urteilen unter Beobachtung. In Stefan Nicolae, Martin Endreß, Oliver Berli, & Daniel Bischur (Hrsg.), *(Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit* (S. 275-298). [https://doi.org/10.1007/978-3-658-21763-1\\_12](https://doi.org/10.1007/978-3-658-21763-1_12).
- Hornbostel, Stefan & Möller, Torger (2015). Die Exzellenzinitiative und das deutsche Wissenschaftssystem: eine bibliometrische Wirkungsanalyse. *Wissenschaftspolitik im Dialog* 12. Berlin: Berlin-Brandenburgische Akademie der Wissenschaften.
- Klawitter, Maren (2015). Effects of Institutional Changes on Requirements for Vacant Professorships in Germany: A Quantitative Content Analysis of Employment Advertisements. *Working Papers in Higher Education Studies*, 1 (2). <http://www.wphes-journal.eu/index.php/wphes/article/view/15>.
- Kleimann, Bernd & Klawitter, Maren (2017a). Berufungsverfahren an deutschen Universitäten aus Sicht organisationaler Akteure. *Beiträge zur Hochschulforschung*, 39 (3/4), 52-73. Verfügbar unter: <http://www.bzh.bayern.de/uploads/media/3-4-2017-Kleimann-Klawitter.pdf>
- Kleimann, Bernd & Klawitter, Maren (2017b). An Analytical Framework for Evaluation-Based Decision-Making Procedures in Universities. In Jeroen Huisman & Malcolm Tight (Eds.), *Theory and method in higher education research* (First edition). Bingley, UK: Emerald Publishing Limited.
- Kleimann, Bernd, In der Smitten, Susanne & Klawitter, Maren (2015). Forschungserfahrung als Berufungskriterium. Anforderungen an zukünftige Professorinnen und Professoren. *Forschung und Lehre*, 2015 (8), 644-645.
- Klein, Tamar, Kraatz, Alexandra & Hornbostel, Stefan (2012). Begutachtungsprozesse im Wettbewerb um Drittmittel. Das Beispiel der Sonderforschungsbereiche. *Die Hochschule*, 2012 (2), 164-182.
- Krüger, Anne K. & Reinhart, Martin (2017). Theories of Valuation-Building Blocks for Conceptualizing Valuation Between Practice and Structure. *Historical Social Research/Historische Sozialforschung*, 263-285. <https://doi.org/10.12759/hsr.42.2017.1.263-285>.
- Krüger, Anne K. & Reinhart, Martin (2016). Wert, Werte und (Be)Wertungen – Ansätze für eine Soziologie des Wertens und Bewertens. *Berliner Journal für Soziologie*. <https://doi.org/10.1007/s11609-017-0330-x>.
- Lamont, Michèle (2012). Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology*, 38, 201-221. <https://doi.org/10.1146/annurev-soc-070308-120022>.
- Luhmann, Niklas (1983). *Legitimation durch Verfahren*. Frankfurt am Main: Suhrkamp.
- Manville, Catriona, Guthrie, Susan, Henham, Marie-Louise, Garrod, Bryn, Sousa, Sonia, Kirtley, Anne, Castle-Clarke, Sophie & Ling, Tom (2015). *Assessing Impact Submissions for REF 2014: An Evaluation*. RAND Europe: Cambridge, UK.
- Meier, Frank, Peetz, Thorsten & Waibel, Désirée (2016). Bewertungskonstellationen. Theoretische Überlegungen zur Soziologie der Bewertung. *Berliner Journal für Soziologie*, 26(3), 307-328. <https://doi.org/10.1007/s11609-017-0325-7>.
- Möller, Torger, Antony, Philipp, Hinze, Sybille & Hornbostel, Stefan (2012). Exzellenz begutachtet. Befragung der Gutachter in der Exzellenzinitiative. *iFQ-Working Paper No.11*. Berlin.

- Neidhardt, Friedhelm (2016). Selbststeuerung der Wissenschaft durch Peer-Review-Verfahren. In Dagmar Simon, Andreas Knie, Stefan Hornbostel, & Karin Zimmermann (Hrsg.), *Handbuch Wissenschaftspolitik* (2. Aufl., S. 261-277). <https://doi.org/10.1007/978-3-658-05455-7>.
- Niewöhner, Jörg (2015). Infrastructures of Society, Anthropology of. *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition (12), 119-125. <https://doi.org/10.18452/19371>.
- Niewöhner, Jörg (2014). Perspektiven der Infrastrukturforschung: Care-ful, relational, ko-laborativ. In Diana Lengersdorf & Matthias Wieser (Hrsg.), *Schlüsselwerke der Science & Technology Studies* (2014. Aufl., S. 341-352). Wiesbaden: VS Verlag.
- Olbrecht, Meike (2014). *Entscheidungsfindungsprozesse von Gutachtergruppen*. Humboldt-Universität zu Berlin, Philosophische Fakultät III. <https://doi.org/10.18452/16946>.
- Pothast, Jörg (2017). The Sociology of Conventions and Testing. In Claudio E. Benzecry, Monika Krause, & I. Ariail Reed (Eds.), *Social Theory Now*. The University of Chicago Press.
- Reinhart, Martin (2012). *Soziologie und Epistemologie des Peer Review*. (1. Aufl.) Baden-Baden: Nomos.
- Rosanvallon, Pierre (2018). *Die gute Regierung* (1. Aufl.; M. Halfbrodt, Übers.). Berlin: Suhrkamp Verlag.
- Schimank, Uwe (2016). Governance der Wissenschaft. In Dagmar Simon, Andreas Knie, Stefan Hornbostel & Karin Zimmermann (Hrsg.), *Handbuch Wissenschaftspolitik* (2. Aufl., S. 39-57).
- Star, Susan Leight (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, 43 (3), 377-391. <https://doi.org/10.1177/00027649921955326>.
- Star, Susan Leight & Bowker, Geoffrey C. (2002). How to Infrastructure. In Leah A. Lievrouw & Sonia Livingstone (Eds.), *Handbook of New Media: Social Shaping and Consequences of ICTs* (pp. 151-162).
- Star, Susan Leight & Ruhleder, Karen (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, 7 (1), 111-134. <https://doi.org/10.1287/isre.7.1.111>.
- Weingart, Peter (2001). *Die Stunde der Wahrheit?: Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft*. Weilerswist: Velbrück Wissenschaft.
- Weller, Ann C. (2001). *Editorial Peer Review: Its Strengths and Weaknesses*. Information Today, Inc.
- Wilholt, Torsten (2012). *Die Freiheit der Forschung: Begründungen und Begrenzungen*. Originalausgabe. Berlin: Suhrkamp Verlag.

### *Anschriften der Autor\*innen:*

Dr. Cornelia Schendzielorz, Humboldt-Universität zu Berlin, Institut für Sozialwissenschaften, Unter den Linden 6, 10099 Berlin, E-Mail: [schendzc@hu-berlin.de](mailto:schendzc@hu-berlin.de) und Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Abteilung Forschungssystem und Wissenschaftsdynamik, E-Mail: [schendzielorz@dzhw.eu](mailto:schendzielorz@dzhw.eu).

Prof. Dr. Martin Reinhart (Korrespondenzautor), Humboldt-Universität zu Berlin, Institut für Sozialwissenschaften, Unter den Linden 6, 10099 Berlin, E-Mail: [martin.reinhart@hu-berlin.de](mailto:martin.reinhart@hu-berlin.de).