

November 2022



Forschungs- evaluation an der Universität Konstanz

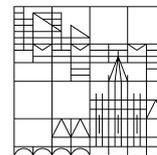


Vorarbeiten, Entwicklungspotentiale und Maßnahmen

Verabschiedet vom Senat der Universität Konstanz am 16.11.2022

Inhalt

1. Einleitung, Ziel und Zweck	3
2. Gegenstände der Forschungsevaluationen	4
2.1. Forschungsrahmenbedingungen	5
2.2. Wissenschaftliche Einrichtungen	5
3. Beurteilungskriterien: Fachspezifische Indikatoren	7
4. Vorgehen und Methodik	7
5. Umsetzung bis 2026	10
Literaturverzeichnis	12



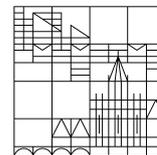
1. Einleitung, Ziel und Zweck

Die Bewertung und Evaluation von Forschungsleistungen sind inhärenter Bestandteil wissenschaftlichen Arbeitens. Sie erfolgen bei der Begutachtung von Publikationen, bei Berufungen, bei der Begutachtung von Qualifikationsarbeiten oder der Vergabe von Forschungsmitteln. In jüngerer Zeit gibt es eine lebhafte Debatte darüber, wie Forschungsleistungen bewertet werden sollten. Inzwischen gibt es eine Vielzahl von Deklarationen, welche zumeist Prinzipien der Leistungsbewertung konstatieren, darunter etwa die San Francisco Declaration on Research Assessment (DORA 2012), das Leiden Manifesto (Hicks, Diana, Wouters, Paul et al. 2015) und zuletzt ein Reformvorschlag der Europäischen Kommission (European Commission - Directorate-General for Research and Innovation 2021). Hierbei wird nicht nur die Frage nach der „richtigen“ Bewertung individueller Forschungsleistungen gestellt, sondern auch die Frage nach der Art von Evaluation ganzer Leistungseinheiten.

Befördert durch institutionelle Förderformate wie die Exzellenzinitiative bzw. –strategie gewinnen auch hochschulinterne Evaluationen von Forschungsstrukturen und -leistungen an Bedeutung. Hochschulleitungen benötigen einen Überblick über die Qualität ihrer Forschungsstrukturen und Forschungsleistungen um Profilbildung zu betreiben und strategische Überlegungen im Hinblick auf neue Forschungsschwerpunkte oder mögliche Clusteranträge vorzunehmen. Auch an der Universität Konstanz ist das Thema Qualitätsmanagement des Leistungsbereichs Forschung präsent: Angestoßen durch den ersten Antrag zur institutionellen Förderung der Exzellenzinitiative zielte die Universität auf den Aufbau eines Qualitätsmanagementsystems (QMS) in allen Leistungsbereichen, dessen Leitlinien und Bausteine im Qualitätsmanagement-Konzept von 2014 beschrieben wurden (siehe *QM-Webseite*). Demnach basiert das QMS der Universität Konstanz „auf Qualitätszielen, klaren Strukturen und Prozessen, kontinuierlichem Monitoring und Evaluationen sowie Strategiegelgesprächen zwischen Rektorat und Fachbereichen.“

Zur Evaluation des Leistungsbereichs Forschung ist es zunächst notwendig zu bestimmen, was an der Universität Konstanz unter Forschungsqualität verstanden wird. Dafür ist die Definition von Zielen, Kriterien und Indikatoren nötig, welche die Qualität von Forschung näher bestimmen und als Beurteilungsmaßstab fungieren (siehe QM-Konzept, S. 7). Eine solche Definition wurde in der Qualitätsmatrix Forschung erarbeitet (verabschiedet vom Senat am 18.11.2018, [Qualitätsmatrix](#)). Sie legt folgende übergeordnete Ziele für die Gesamtuniversität fest: (i) Attraktivität für herausragende Wissenschaftler*innen, (ii) Exzellente Forschung von exzellenten Wissenschaftler*innen und (iii) Optimale Rahmenbedingungen. Seit 2012 werden bereits die Forschungsleistungen der Fachbereiche im Rahmen sogenannter Peer-Review-Zyklen durch externe Gutachter*innen bewertet.

Um das Thema Forschungsevaluation vertieft zu betrachten und die Frage nach angemessenen Kriterien, validen Indikatoren und effektiven Verfahren zu stellen, wurde im Rahmen der Exzellenzstrategie im Jahr 2019 das Exzellenzprojekt „Qualitätsmanagement in Forschung und Nachwuchsförderung“ etabliert. Das Projekt steht dabei in engem Zusammenhang mit zwei weiteren Exzellenzprojekten - dem FIS-Projekt, welches den Aufbau eines Forschungs-



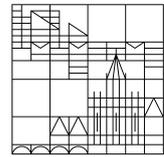
informationssystem (FIS) koordiniert, und dem Projekt BICC (Business Intelligence Competence Centre), welches über Forschungsinformationen die datenbasierte Steuerung unterstützt - und baut auf diesen auf. Von regelmäßigem Austausch und der Nutzung von Synergien bis hin zu operativer Unterstützung entsteht hier ein sich wechselseitig befruchtendes Umfeld im Bereich Forschungsinformation, -controlling und -evaluation.

Das vorliegende Papier ist ein Ergebnis des Projekts „Qualitätsmanagement in Forschung und Nachwuchsförderung“. Es zielt darauf, Forschungsevaluationen an der Universität Konstanz zukünftig methodisch fundierter durchzuführen und den gesamten Evaluationskreislauf - von der Festlegung der Evaluationsgegenstände über die Definition von Beurteilungskriterien bis hin zu Vorgehen und Methodik – klarer zu fassen, um die Nützlichkeit der Evaluationen zu erhöhen.

Die *Gegenstände der Forschungsevaluationen* (Kap. 2) sollen erweitert werden: Neben der Attraktivität der Universität und den Forschungsleistungen, sollen die *Rahmenbedingungen* der Forschung besser abgebildet werden (Kap. 2.1). Bei der Evaluation von *wissenschaftlichen Einrichtungen* (Exzellenzcluster, Zukunftskolleg) zeigen sich neue Anforderungen (bspw. ein stärkerer Fokus auf Rahmenbedingungen) und eine steigende Nachfrage, welche eine Standardisierung der Evaluationen sinnvoll erscheinen lassen (siehe Kap. 2.2). Als Grundlage für Evaluationen sollen zudem die *Beurteilungskriterien* zusammen mit den Fachbereichen um fachspezifische Indikatoren erweitert werden, um die unterschiedlichen Fachkulturen ausreichend zu berücksichtigen (Kap. 3). Bei *Vorgehen und Methodik* bedarf es darüber hinaus zusätzlicher Methoden für die Evaluation von Forschungsleistungen (Kap. 4). Für aussagekräftige, qualitative Gutachten in der Forschung mangelt es an systematischen qualitativen Informationen zu inhaltlichen Forschungsschwerpunkten und zu den wichtigsten Arbeitsergebnissen der Fachbereiche. Diese sollen durch sinnvolle quantitative Indikatoren in Absprache mit den Fachbereichen und im Einklang mit bibliometrischen Erkenntnissen ergänzt werden. Kapitel 5 fasst abschließend zusammen, anhand welcher Maßnahmen die Desiderate angegangen werden sollen (siehe auch Tabelle 1).

2. Gegenstände der Forschungsevaluationen

Als Grundlage für Forschungsevaluationen an der Universität Konstanz dient die Qualitätsmatrix Forschung. Sie rückt als Evaluationsgegenstände insbesondere die Attraktivität der Universität/der Fachbereiche, die Forschungsleistungen und die Forschungsrahmenbedingungen in den Fokus. Aufgrund mangelnder Informationen im Bereich Forschungsrahmenbedingungen wurden in bisherigen Forschungsevaluationen vor allem Attraktivität und Forschungsleistungen der Fachbereiche evaluiert (siehe auch Vorgehen und Methodik, Kap. 4), kaum jedoch die Forschungsrahmenbedingungen. Daher soll die Informationsgrundlage für die Gutachtenden in diesem Bereich insbesondere durch Befragungen ausgebaut werden (siehe Kapitel 2.1). Zudem sind immer häufiger nicht nur Fachbereiche, sondern zentrale, wissenschaftlichen Einrichtungen Gegenstand der Evaluation. Die Evaluationen wissenschaftlicher Einrichtungen sollen daher standardisiert werden, um die Aufwände aller Beteiligten zu minimieren (siehe Kapitel 2.2).



2.1. Forschungsrahmenbedingungen

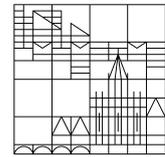
Die Qualitätsmatrix Forschung nennt als ein Qualitätsziel der Universität Konstanz „Exzellente Rahmenbedingungen“ für Wissenschaftler*innen und Forschung. Bisher wurden jedoch nur sehr vereinzelt Informationen zur Erreichung dieses Ziels erhoben oder bereitgestellt.

Vorarbeiten – Im Rahmen der Promovierenden- und der Promoviertenbefragungen wurden in regelmäßigen Abständen Informationen etwa zu den Promotionsbedingungen oder zur Finanzierung der Promovenden aufbereitet. In den wiederkehrenden Peer-Reviews wurden auch schon auf Wunsch der Fachbereiche Befragungen zur Zufriedenheit der Mitarbeitenden durchgeführt. Im Rahmen der Kund*innenbefragung zur Evaluation zentraler wissenschaftsunterstützender Einrichtungen und zur Evaluation der Exzellenzstrategie wurden ebenfalls Fragen zur Zufriedenheit der Wissenschaftler*innen gestellt. Zu nennen sind in diesem Zusammenhang auch die Qualitätssicherungsmaßnahmen für Forschungsinfrastrukturen durch die Stabsstelle Universitätsentwicklung, Forschung und Transfer.

Entwicklungspotential und Maßnahmen – Um die Frage beantworten zu können, inwieweit das Qualitätsziel „Optimale Rahmenbedingungen“ erreicht wird, ist geplant die bisherigen Befragungen durch eine gesonderte Befragung zu Rahmenbedingungen mit einem 5-jährigen Befragungszyklus zu ergänzen und bereits vorhandene Befragungen, um passende Frageblöcke zu erweitern. Hierbei orientieren wir uns an den Kriterien der Qualitätsmatrix. Das Kriterium „Unterstützende Strukturen und Dienste“ soll die „Nutzung und Zufriedenheit mit unterstützenden Strukturen und Diensten und Forschungsbedingungen allgemein“ und den „(z)eitliche(n) Aufwand für Administration / akad. Selbstverwaltung / Lehre (im Verhältnis zur Forschung)“ abfragen. Darüber hinaus wird ein „produktives Umfeld“ als Desiderat in der Qualitätsmatrix genannt, welches abgedeckt werden sollte. Fragen, um die Produktivität des Umfelds zu messen, sollten miteinbeziehen, inwiefern sich Fachbereichsmitglieder inklusiv behandelt fühlen, ob fachbereichsweite Strategien partizipativ entwickelt und gut kommuniziert werden und inwiefern eine Vernetzung der Fachbereichsmitglieder mit den Forschungs- und Nutzergemeinschaften sichergestellt wird. Beim Kriterium „Finanzielle und personelle Ausstattung“ sollte die Bewertung interner Mittelvergabeprozesse und Förderprogramme im Zentrum stehen und zum Kriterium „Infrastrukturausstattung“ sollten Informationen aus bestehenden Nutzungsanalysen, ggf. durch zusätzliche Zufriedenheitsumfragen, auch unter externen Nutzern, ergänzt werden. Dabei sollten bspw. auch die Themen Open Science und Prozesse zur Einhaltung ethischer Standards mitgedacht werden. Als Erhebungsmethoden eignen sich zur Erfassung oben genannter Datenbedarfe neben Befragungen auch Dokumentenanalysen.

2.2. Wissenschaftliche Einrichtungen

Im Gegensatz zu den wiederkehrenden Peer-Review-Zyklen der Fachbereiche fanden Evaluationen zentraler wissenschaftlicher Einrichtungen bisher vor allem anlassbezogen statt. Dieses Vorgehen ermöglicht die Entwicklung einer passgenauen und an die Bedarfe der einzelnen Einrichtung abgestimmten Evaluationskonzeption, ist jedoch vergleichsweise aufwändig und ressourcenintensiv. Angesichts des steigenden Bedarfs soll dieses Vorgehen durch eine

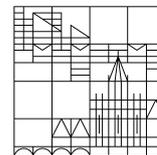


standardisierte Vorgehensweise ersetzt werden, was wiederum eine vertiefte Auseinandersetzung mit allgemeinen Kriterien und Vorgehensweisen für die Evaluation wissenschaftlicher Einrichtungen bedingt.

Vorarbeiten – Im Rahmen des Exzellenzclusters „Kulturelle Grundlagen von Integration“ wurde eine Zwischenevaluation durchgeführt. Zudem wurde das Zukunftskolleg bereits dreimal durch die Stabsstelle QM evaluiert (siehe QM-Webseite für weitere Evaluationen zentraler wissenschaftlicher Einrichtungen). Demnächst stehen die Evaluationen folgender wissenschaftlicher Einrichtungen an: Exzellenzcluster In_equality, Exzellenzcluster Collective Behaviour und Zentrum für kulturwissenschaftliche Forschung. Auch das in der Gründung befindliche Binationale Zentrum Frühe Kindheit hat eine vierjährige Evaluation in seine Satzung geschrieben. Das bisher angewandte Vorgehen bei anlassbezogenen Evaluationen lässt sich wie folgt beschreiben: Zuerst werden für jeden Einzelfall Ziel und Zweck, Gegenstand und Beteiligte der Evaluation festgelegt. Zusammen mit den Beteiligten wird in einem nächsten Schritt zum Evaluationsgegenstand ein Beurteilungsraster mit Zielen, Kriterien und Indikatoren entworfen. Im Anschluss findet die Erhebung der Indikatoren statt. Mögliche Erhebungsmethoden sind Experteninterviews, Dokumentenanalysen, Befragungen und hochschulstatistische Kenngrößen sowie Begehungen durch Peers. Die Interpretation der Daten erfolgt in einem Evaluationsbericht bzw. Gutachter*innenbericht, aus welchem dann Entwicklungsziele und ggf. Maßnahmen abgeleitet werden.¹

Entwicklungspotenziale und Maßnahmen – Das oben beschriebene Vorgehen ist einzelfallbezogen und daher mit hohem Aufwand verbunden. Daher ist es sinnvoll Evaluationen wissenschaftlicher Einrichtungen zu standardisieren, was den Aufwand in der Konzeption der Evaluationen für alle Beteiligten reduziert und auch die Übertragbarkeit von den Evaluationsergebnissen (etwa von Entwicklungszielen und Konsequenzen) erhöht. In Anlehnung an das Qualitätsziel „Exzellente Rahmenbedingungen“ aus der Qualitätsmatrix Forschung soll ein Beurteilungsraster entwickelt werden, welches bspw. die Zielerreichung und die Rahmenbedingungen der Forschung für wissenschaftliche Einrichtungen berücksichtigt. In Bezug auf die Erhebungsmethoden lassen sich auf der Basis des Beurteilungsrasters dann Leitfragen für Interviews, Standards für die Dokumentenanalysen und sonstiger zu erhebender Informationen entwickeln. Für jede Evaluation kann bei Bedarf ein Schwerpunkt festgelegt werden, welcher aber nur ein Gegenstand von vielen, d.h. nicht der alleinige Gegenstand, der Evaluation wäre.

¹ Beispielsweise war bei der Evaluation des Exzellenzclusters „Kulturelle Grundlagen von Integration“ das Ziel die Zwischenevaluation des Clusters hinsichtlich selbst gesteckter Ziele, um daraus Empfehlungen für zukünftige Exzellenzcluster abzuleiten. Gegenstand der Evaluation war die „Strukturen und Entscheidungsmechanismen“ des Clusters. Beteiligte waren unter anderem Rektorat, Sprecher und Geschäftsführer des Clusters und beteiligte Wissenschaftler*innen. Vor der Evaluation wurden Kriterien entwickelt, welche festlegten wie der Evaluationsgegenstand (Strukturen und Entscheidungsmechanismen) ausgestaltet sein sollten (bspw. Zielerreichung, Reichweite, Transparenz und Akzeptanz, Schnelligkeit, etc.), und dazugehörige Indikatoren entwickelt



3. Beurteilungskriterien: Fachspezifische Indikatoren

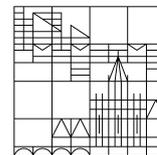
Aus Gründen der Transparenz und Nachvollziehbarkeit ist es sinnvoll a priori Ziele, Kriterien und Indikatoren des zu bewertenden Gegenstands (sei es die Forschungsleistung eines Fachbereichs, eines Forschungsschwerpunkts, individueller Forschenden oder einer wissenschaftlichen Einrichtung) zu definieren. So können externe Gutachtende nachvollziehbar feststellen, ob eine zu evaluierende Einrichtung ihre Ziele erreicht hat und/oder wie sich die Einrichtung weiterentwickeln könnte.

Vorarbeiten – Die 2018 vom Senat verabschiedete Qualitätsmatrix Forschung der Universität Konstanz orientiert sich unter anderem am Kerndatensatz Forschung und entstand im Dialog mit Sektionen und Fachbereichen. Die dort festgelegten Qualitätsziele sind jeweils mit Kriterien und dazugehörigen Indikatoren näher spezifiziert. Allerdings beziehen sich diese Indikatoren auf die Gesamtuniversität und können daher nicht die Besonderheiten der einzelnen Fachbereiche oder einzelner Fachgruppen innerhalb der Fachbereiche berücksichtigen. Um dennoch den Bedarfen der Fachbereiche zu genügen, wurde die Informationsgrundlage im Bereich Forschung bei bisherigen Evaluationen der Fachbereiche stets in enger Abstimmung mit den Fachbereichen und dem Controlling zusammengestellt. Um diese Abstimmungsbedarfe für die Zukunft zu reduzieren und transparent eine geeignete Indikatorik für die Fachbereiche zu ermitteln, wurden vom Rektorat auf der Ebene der Fachbereiche Diskussionen zu fachbereichsspezifischen Qualitätsindikatoren angestoßen. Die Indikatoren sollen in sogenannten Fachbereichsqualitätsmatrizen verabschiedet werden, welche auf die in der universitätsweiten Qualitätsmatrix Forschung genannten übergeordneten Ziele aufbauen. Inzwischen liegen bereits die Qualitätsmatrizen mehrerer Fachbereiche vor.

Entwicklungspotential und Maßnahmen – Auf Basis der Qualitätsmatrizen der Fachbereiche, welche voraussichtlich bis Mitte 2022 vorliegen, werden in einem nächsten Schritt die Zusammenstellung der Forschungsinformationen für die wiederkehrenden Monitorings der Fachbereiche (Peer-Review-Zyklus und Rektoratszyklus) überarbeitet. Hierbei eventuell neu entstehende Datenbedarfe werden zusammen mit den BICC- und FIS-Projekten besprochen und auf technische Umsetzbarkeit hin geprüft.

4. Vorgehen und Methodik

Wie bereits in der Einleitung ausgeführt, wird die Praxis der Bewertung von Forschungsleistungen anhand von Kennzahlen seit geraumer Zeit und aktuell intensiv diskutiert. Daher soll im Folgenden das Vorgehen und die Methodik zur Bewertung von Forschungsleistungen besprochen werden. In anderen Bereichen, wie etwa der Bewertung von Rahmenbedingungen oder der Attraktivität der Fachbereiche kann auf etablierte Methoden aus anderen Evaluationsbereichen zurückgegriffen werden (bspw. Befragungen und hochschulstatistische Größen). Einfache Kennzahlen wie der Umfang der eingeworbenen Drittmittel oder die Anzahl der Publikationen eines Fachbereichs sind zwar als orientierende Begleitinformationen nützlich; sie sollen und können jedoch nicht inhaltliche Schwerpunktsetzungen oder gar die Qualität der erarbeiteten wissenschaftlichen Inhalte abbilden. Die Evaluationsstandards der Leibniz-Gemeinschaft, der DFG und anderer Drittmittelgeber*innen sehen daher zumeist die qualitative



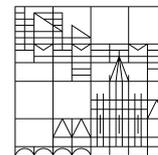
Bewertung konkreter Arbeitsergebnisse durch informed peers vor. Darüber hinaus sollen die präsentierten quantitativen Kennzahlen auf ihre bibliometrische Validität hin überprüft werden. Insgesamt soll der mixed-methods approach der Stabsstelle QM auch im Bereich Forschung angewandt und verfeinert werden.

Vorarbeiten – Bisher wird in den Verfahren ein Fokus auf einen „informed peer-review“ gelegt. Das heißt, dass in Monitoringbericht und Selbstbeurteilung im Monitoringbericht den „peers“ systematisch Informationen zur Anzahl der (Peer-Review)-Publikationen sowie zu Anzahl und Summen der Drittmittelprojekte nach Geldgebern präsentiert wurden. Auch wichtige Verbundprojekte, Preise, Infrastrukturen, Kooperationen und Transferleistungen wurden aufbereitet. Ein Anhang listete darüber hinaus die Forschungsprojekte der Fachbereiche auf. Die präsentierten Informationen wurden bisher in enger Abstimmung mit den Fachbereichen zusammengestellt. Zudem wurden die Fachbereiche zur Verabschiedung eigener Qualitätsmatrizen aufgefordert (siehe Kap. 3). Damit wird der Abstimmungsaufwand zukünftig reduziert: Anhand der Qualitätsmatrizen der Fachbereiche sollen die oben genannten bereits vorhandenen Kennzahlen angepasst und ggf. ergänzt werden, sofern die Informationen verfügbar sind.

Entwicklungspotential und Maßnahmen – Zwei komplementäre Entwicklungspotentiale müssen in den Bereichen quantitative Kennzahlen und qualitativen Informationen angegangen werden:

Quantitative Informationen – Im engeren Sinne bezieht sich ein Großteil der Kritik an der Messung von Forschungsleistungen auf die Praxis, die Qualität einer Publikation durch den Impact Factor oder das Prestige der Zeitschrift, in welcher sie veröffentlicht wurde, zu approximieren (siehe bspw. (DORA 2012)). In Berufungsverfahren beispielsweise könnten Gutachtende eine Vorauswahl der Bewerber*innen anhand der Impact Faktoren treffen, bspw. durch Addition der Impact Factors der jeweiligen Publikationen je Wissenschaftler*in. Die Verwendung des Impact Factors als Indikator für die wissenschaftliche Signifikanz einer Publikation zeigt beispielsweise, dass eine bibliometrische Fundierung in diesem Bereich wichtig ist (Moher et al. 2018; Engwall 2014; Bornmann und Marx 2014). Die Veröffentlichung in einem Journal mit hohem Impact Factor sagt jedoch hauptsächlich etwas über das von Editoren und Gutachtenden vermutete wissenschaftliche Potential aus; nicht jedoch über dessen tatsächliche Signifikanz.² Neuere Ansätze versuchen beispielsweise die starken zeit- und feldabhängigen Variationen in den Zitationen durch die Kalkulation von Perzentilen je Publikationsjahr und

² Grob gesagt wird unter dem Impact Factor einer Zeitschrift zumeist die Anzahl aller Zitationen aller Artikel in einem Jahr, geteilt durch Anzahl der zitierfähigen Artikel, welche in den vorigen zwei Jahren in der Zeitschrift erschienen sind. Oft wird eine Publikation in einer hochrangigen Zeitschrift per se als Auszeichnung gewertet. Dies mag unter anderem daran liegen, dass der wissenschaftliche Erfolg eines Beitrags selbst als stochastischer Prozess identifiziert wird: Retrospektiv stellt sich oft heraus, dass die Signifikanz wichtiger Beiträge erst einige Jahre nach ihrer Veröffentlichung entdeckt und auch erst dann entsprechend zitiert wird. In der Tat ist die Verteilung von Zitationen der Artikel einer Zeitschrift meist schief: Einige wenige Artikel werden sehr häufig zitiert; der Großteil jedoch selten oder gar überhaupt nicht (Albarrán et al. 2011). Als Folge dessen sagt der Impact Factor einer Zeitschrift nur sehr bedingt etwas über die Zitationshäufigkeit eines zu bewertenden Beitrags aus. Der 2-Jahres-Impact Factor (definiert nach Web of Science/Thomson Reuters) kann – wie üblich beim arithmetischen Mittel – aufgrund einzelner „Ausreißer“ (vielzitierte Artikel) im Zeitverlauf stark variieren (Engwall 2014). Auch unterscheiden sich beispielsweise die Qualität der Indizierung, Publikationstypen und Zitationsgewohnheiten als auch Verbreitungsgeschwindigkeiten teilweise stark zwischen Disziplinen und innerhalb von Subdisziplinen (Fuchs 2014). Daher wird oft kritisiert, dass beispielsweise das gewählte Zeitfenster zu kurz ist (van Leeuwen 2012) oder die zu Grunde gelegten Publikations-

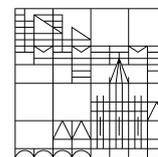


wissenschaftlichen Feld abzumildern und weisen den Anteil der Publikationen im oberen Dezil je Publikationsjahr und Feld als Erfolgsindikator aus (Bornmann und Leydesdorff 2014). Auch der h-index steht in der Kritik; nicht zuletzt weil er die für die Forschung verfügbare Zeit, das akademische Alter (Bonaccorsi und Daraio 2003), nicht berücksichtigt und dadurch Nachwuchswissenschaftler*innen tendenziell benachteiligt. Hier gilt es für die Bewertung von Forschungsleistungen bibliometrisch sinnvolle Praktiken im Sinne von „responsible metrics“ aufzuzeigen (James Wilsdon et al. 2015) und diese insbesondere in die Evaluationen der Stabsstelle QM zu integrieren. Oft bedeutet das, eine Vielzahl von Kennzahlen auszuweisen und teilweise wird es auch bedeuten gewisse Kennzahlen, etwa zu Zitationen, aufgrund mangelnder Datenqualität in bestimmten Forschungsfeldern nicht zu berichten. Hier sollte ein Verständigungsprozess (und ggf. auch eine Beratung/Sensibilisierung) darüber stattfinden, was geeignete Kennzahlen sind, die auch das messen, was intendiert ist.

Qualitative Informationen – In verschiedenen Kontexten wurde gezeigt, dass eine Überbewertung von einfachen Indikatoren wie der Anzahl der Publikationen oder dem Drittmittelaufkommen zu einem geringeren Fokus auf hochqualitative Forschung führen kann, was in einem Rückgang bestimmter Publikationstypen, Rückgang absoluter Zitationszahlen und einer Zunahme wissenschaftlichem Fehlverhaltens gesehen werden kann (Hammarfelt und Rijcke 2015; Ingwersen und Larsen 2014; Council of Canadian Academies 2012; Butler 2003; James Wilsdon et al. 2015; Engwall 2014). In jüngerer Zeit gab es daher Anstrengungen diesem Trend entgegenzuwirken und die qualitative Beurteilung einzelner Arbeitsergebnisse durch Wissenschaftler*innen in den Fokus zu rücken (Europäische Kommission 2021; Kleiner 2017; Research Excellence Framework (UK)). Auch in den Monitoringverfahren der Stabsstelle QM sollen qualitative Beurteilungsmethoden stärker in den Prozess miteinbezogen werden.³ Passend hierzu stellen wir fest, dass in vielen Fachbereichsqualitätsmatrizen der Fachbereiche ein Indikator wie „Auswahl der 10 wichtigsten Publikationen ...“ vorgeschlagen wurde, um das Qualitätskriterium für „Quantität und Qualität des Outputs“ zum übergeordneten Ziel „Exzellente Forschung“ abzubilden. Interessanterweise sehen auch die Grundsätze für Forschungsevaluation wichtiger Geldgeber (bspw. der DFG oder der Leibniz Gemeinschaft) vor, dass eine zu evaluierende Einrichtung einige wichtige Arbeitsergebnisse auswählt. Wie in den Leibniz-Evaluationsverfahren (Leibniz Gemeinschaft 2018) soll die Praktikabilität dabei über eine kurze Beschreibung der drei bis zehn wichtigsten Arbeitsergebnisse des Fachbereichs sicher-

typen (in der Hauptsache englischsprachige Zeitschriftenartikel) zu eng gefasst sind (van Raan et al. 2011). Erweiterungen oder Änderungen des Impact Factors, die die oben genannten Problematiken abfedern sollen, sie jedoch nicht grundlegend entkräften, sind der 5-Jahres-Impact-Factor, Source-Normalized Impact Factor (welcher Zitationen nach Impact Factor des zitierenden Journals und Feld gewichtet mit Datenquelle Scopus/Elsevier) oder die Normalized Journal Position/Q1-Journals (welcher den Rang innerhalb eines wiss. Feldes/das obere Quartil anhand des Impact Factors misst, um Vergleichbarkeit zwischen Feldern herzustellen).

³ Qualitativ ist hier neutral im Sinne von die Summe aller Eigenschaften betreffend zu verstehen, auch wenn solche qualitativen Informationen zusätzlich die Bewertung der Güte der Eigenschaften im wertenden Sinne von „qualitativ“ erleichtern: Nach den Leibniz-Grundsätzen wird die Forschungsqualität nach den Kriterien Qualität, Resonanz, Rezeption, Wirkung, Wettbewerbsfähigkeit, Sichtbarkeit, Quantität, Originalität, Aktualität, Risikobereitschaft, Kohärenz, Schlüssigkeit und Relevanz für andere gesellschaftliche Bereiche bewertet. Teilbereiche werden abschließend als exzellent, sehr gut, etc. bewertet. Das Research Excellence Framework (UK) sieht vor jedes Arbeitsergebnis nach fachspezifisch definierten Kriterien (originality, significance und rigour) als world-leading, internationally excellent, recognised internationally, oder recognised nationally zu bewerten.



gestellt werden, ggf. gegliedert nach inhaltlichen Schwerpunkten. Arbeitsergebnisse sollen dabei in Einklang mit einer kürzlich eingeführten Innovation der Leibniz-Grundsätze in einem weiteren Sinne gefasst werden: Sie können Transferleistungen (wie Patente oder Ausstellungen) oder die Bereitstellung von Infrastruktur (bspw. einer Online-Datenbank oder eines Großgeräts) umfassen, aber auch (noch) nicht veröffentlichte Zwischenergebnisse wie Datenpublikationen. Diese Arbeitsergebnisse sollen dann von der externen Gutachtergruppe qualitativ beurteilt werden.

Darüber hinaus erlaubt die Informationsgrundlage in den Peer-Reviews bisher keine systematische Verortung der inhaltlichen Schwerpunkte der Fachbereiche. Die inhaltlichen Schwerpunkte sind jedoch wichtig um die Beurteilung von Kooperationspotentialen zwischen den Disziplinen innerhalb der Universität aber auch mit externen Kooperationspartnern zu vereinfachen. Um dieses Entwicklungspotential systematisch anzugehen, soll eine Restrukturierung der Leitfragen für die Peer-Reviews der Fachbereiche und ergänzend statistische, semantische Analysen von Publikationsabstracts durchgeführt werden, was beispielsweise in Zusammenarbeit mit dem BICC-Projekt geschehen könnte.

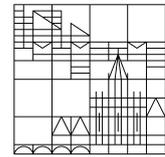
Ausblick – Beide Ansätze, quantitativ und qualitativ, sollen verfolgt werden, um ein kompletteres Vorgehen für die Evaluationen zu entwickeln. Das so entstandene Wissen kann – bei Bedarf – auch für Fachbereiche hilfreich sein. Auch bei der Entwicklung einer Position der Universität Konstanz zu allgemeinen Fragen der Forschungsbewertung könnten die Erfahrungen aus den Evaluationen hilfreich sein, etwa bei der Entwicklung von Richtlinien für offene Fragen der Forschungsbewertung (bspw. in Form einer „Research Assessment Policy“).

5. Umsetzung bis 2026

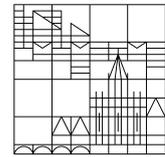
Insgesamt soll der Bereich Forschungsevaluation den Gegenstand Rahmenbedingungen stärker in den Fokus nehmen. Die Evaluation wissenschaftlicher Einrichtungen soll vereinheitlicht werden. Auf Basis der fachbereichsweit geteilten Qualitätsvorstellungen soll die Bewertung von Forschungsleistungen verbessert werden: Zum einen durch die Anreicherung der Evaluationen um qualitative Informationen, zum anderen durch die Unterstützung der qualitativen Informationen durch geeignete, d.h. bibliometrisch fundierte, Kennzahlen. In Zusammenarbeit mit den Fachbereichen und wissenschaftlichen Einrichtungen möchten wir damit unseren Beitrag zum Erhalt des vitalen und kreativen Forschungsumfelds der Universität Konstanz leisten. Tabelle 1 liefert einen Überblick der geplanten Maßnahmen.

Tabelle 1 Maßnahmenübersicht

Desiderat	Maßnahme	Zuständigkeit	Zeitraumen
Erfassung der Rahmenbedingungen der Forschung	Befragungskonzeption zu Rahmenbedingungen der Forschung	QM, Stabsstelle Universitätsentwicklung, Forschung	bis 2023-24



		und Transfer (FIS, Forschungssupport), Fachbereiche	
Evaluation wissenschaftlicher Einrichtungen	Beurteilungsraster und standardisiertes Evaluationsverfahren für wissenschaftliche Einrichtungen entwickeln	QM, zentrale wissenschaftl. Einrichtungen	bis 2023-25
Fachspezifische Qualitätsvorstellungen	Sobald die Fachbereichsqualitätsmatrizen durch die FB verabschiedet: Aggregation der Informationsbedarfe für eigene Verfahren, FIS und Controlling (BICC) und Überarbeitung der Monitoringberichte	Fachbereiche, QM, Controlling (BICC), Stabsstelle Universitätsentwicklung, Forschung und Transfer (FIS)	ab 2023 und je nach Datenverfügbarkeit
Vorgehen und Methodik	Erweiterung von Verfahrensunterlagen / der bereitgestellten Informationen um qualitative Informationen in Anlehnung an die Leibniz-Grundsätze und wissenschaftlich fundierte quantitative Informationen	QM, Fachbereiche	Kontinuierlich und in den anstehenden Peer-Reviews



Literaturverzeichnis

Albarrán, Pedro; Crespo, Juan A.; Ortuño, Ignacio; Ruiz-Castillo, Javier (2011): The skewness of science in 219 sub-fields and a number of aggregates. In: *Scientometrics* 88 (2), S. 385–397. DOI: 10.1007/s11192-011-0407-9.

Bonaccorsi, Andrea; Daraio, Cinzia (2003): Age effects in scientific productivity. In: *Scientometrics* 58 (1), S. 49–90.

Bornmann, Lutz; Leydesdorff, Loet (2014): Scientometrics in a changing research landscape. In: *EMBO rep* 15 (12), S. 1228–1232. DOI: 10.15252/embr.201439608.

Bornmann, Lutz; Marx, Werner (2014): How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. In: *Scientometrics* 98 (1), S. 487–509. DOI: 10.1007/s11192-013-1161-y.

Butler, Linda (2003): Explaining Australia's increased share of ISI publications—the effects of a funding formula based on publication counts. In: *Research Policy* (32).

Council of Canadian Academies (2012): Informing Research Choices: Indicators and Judgment.

DORA (2012): San Francisco Declaration on Research Assessment. Hg. v. American Society for Cell Biology. Online verfügbar unter <https://sfdora.org/read/>.

Engwall, Lars (2014): On the quality of quality assessments. In: *Use and Abuse in the Review of Research Performance*.

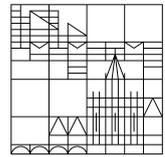
European Commission - Directorate-General for Research and Innovation (2021): Towards a reform of the research assessment system.

Fuchs, Milena Zic (2014): Bibliometrics: use and abuse in the humanities. In: *Use and Abuse in the Review of Research Performance*.

Hammarfelt, B.; Rijcke, S. de (2015): Accountability in context: effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University. In: *Research Evaluation* 24 (1), S. 63–77. DOI: 10.1093/reseval/rvu029.

Hicks, Diana, Wouters, Paul; Waltmann, Ludo; Rijcke, Sarah de (2015): Comment: The Leiden Manifesto for Research Metrics. In: *Nature* (520).

Ingwesen, Peter; Larsen, Birger (2014): Influence of a performance indicator on Danish research production and citation impact 2000–12. In: *Scientometrics* 101 (2), S. 1325–1344. DOI: 10.1007/s11192-014-1291-x.



James Wilsdon; Liz Allen; Eleonora Belfiore; Philip Campbell; Stephen Curry; Steven Hill et al. (2015): *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*.

Kleiner, Matthias (2017): *Selbstorganisation als Erfolgprinzip. Governance, Performance, and Leadership of Reserach and Public Organizations*, 2017.

Leibniz Gemeinschaft (2018): *Grundsätze Evaluierungsverfahren des Senats der Leibniz-Gemeinschaft, Anlage 3 - Muster zur Erstellung einer Evaluierungsunterlage*.

Moher, David; Naudet, Florian; Cristea, Ioana A.; Miedema, Frank; Ioannidis, John P. A.; Goodman, Steven N. (2018): *Assessing scientists for hiring, promotion, and tenure*. In: *PLoS biology* 16 (3), e2004089. DOI: 10.1371/journal.pbio.2004089.

Research Excellence Framework (UK): *Panel criteria and working methods 2019*.

van Leeuwen, Thed (2012): *Discussing some basic critique on Journal Impact Factors: revision of earlier comments*. In: *Scientometrics* 92 (2), S. 443–455. DOI: 10.1007/s11192-012-0677-x.

van Raan, Anthony F. J.; van Leeuwen, Thed N.; Visser, Martijn S. (2011): *Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings*. In: *Scientometrics* 88 (2), S. 495–498. DOI: 10.1007/s11192-011-0382-1.