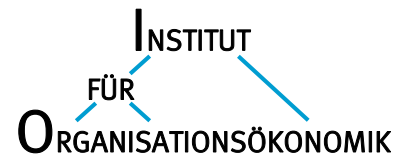




WESTFÄLISCHE  
WILHELMS-UNIVERSITÄT  
MÜNSTER



Diskussionspapier des  
Instituts für Organisationsökonomik

1/2012

Die Zitationshäufigkeit als Qualitätsindikator  
im Rahmen der Forschungsleistungsmessung

Harry Müller

Discussion Paper of the  
Institute for Organisational Economics

**Diskussionspapier des  
Instituts für Organisationsökonomik  
1/2012**

Januar 2012

ISSN 2191-2475

**Die Zitationshäufigkeit als Qualitätsindikator  
im Rahmen der Forschungsleistungsmessung**

*Harry Müller*

**Zusammenfassung**

Dieser Beitrag widmet sich den Problemen und Anwendungsfeldern der Forschungsleistungsmessung mit Hilfe von Zitationen. Dazu ist zunächst grundlegend zu klären, weshalb Zitationen als Indikator für Forschungsleistung geeignet sind und wo die Grenzen dieses Konzepts liegen. Darauf aufbauend sollen die gängigen Methoden zur Erstellung wissenschaftlicher Rankings vorgestellt und kritisch bewertet werden. Dabei stellt sich heraus, dass für eine deutschsprachige Sozialwissenschaft wie die Betriebswirtschaftslehre, an deren Beispiel in diesem Beitrag argumentiert wird, zahlreiche Besonderheiten zu berücksichtigen sind. Bei der Wahl der geeigneten Rankingmethode muss bedacht werden, dass die Publikationen z. T. in deutscher Sprache erscheinen und neben wissenschaftlichen Zeitschriften auch Monographien und Sammelbände wichtige Publikationsmedien sind. Bei Berücksichtigung dessen lässt sich folgern, dass zitationsbasierte Verfahren trotz ihrer spezifischen Schwächen die für viele Zwecke am besten geeignete Methodik darstellen.

JEL-Codes: I23, I20, A11

# **The Citation Frequency as an Indicator for Quality in the Assessment of Research Performance**

## **Abstract**

This article addresses the problems and application fields of using citations as a measure of research performance. First, it is pointed out why citations are a useful indicator of research performance and what defines the limits of this approach. Thereon, the current methods for constructing research rankings will be presented and critically evaluated. It turns out that for a social science like business administration, several peculiarities have to be considered: Research findings are often published in German language, and besides scientific journals, monographs and edited volumes are equally important types of research media. By taking this into account, it can be argued that despite some specific weaknesses citation based approaches are the most appropriate methodology for accessing research performance under many circumstances.

Im Internet unter:

[http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO\\_1\\_2012.pdf](http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO_1_2012.pdf)

Westfälische Wilhelms-Universität Münster  
Institut für Organisationsökonomik  
Scharnhorststraße 100  
D-48151 Münster

Tel: +49-251/83-24303 (Sekretariat)  
E-Mail: [io@uni-muenster.de](mailto:io@uni-muenster.de)  
Internet: [www.wiwi.uni-muenster.de/io](http://www.wiwi.uni-muenster.de/io)

# Die Zitationshäufigkeit als Qualitätsindikator im Rahmen der Forschungsleistungsmessung\*

## 1. Einleitung

Die Definition und die Operationalisierung von Zielgrößen werden im Rahmen des Hochschulmanagements zunehmend wichtiger. Ein zweckmäßiges Management setzt voraus, dass Outputgrößen bekannt sind, die dann unter Berücksichtigung der zu ihrer Erstellung nötigen Ressourcen bewertet und verglichen werden können. Fehlt die Information über den Output, kann weder die Effektivität einer Aktivität bewertet, noch das Effizienzziel sinnvoll angestrebt werden. Es ist dabei eine Besonderheit des Hochschulwesens, dass sowohl Forschung als auch Lehre komplexe Dienstleistungen darstellen, deren Output wesentlich schwieriger zu erfassen ist als die bei ihrer Erbringung verwendeten Ressourcen. Seit etwa zwei Jahrzehnten besteht seitens der Politik als Träger der (öffentlichen) Hochschulen, aber zunehmend auch seitens der Studierenden bzw. der Gesellschaft generell ein gesteigertes Interesse an einer Quantifizierung. Die Exzellenzinitiative des Bundes, Zielvereinbarungen bei der Budgetierung, leistungsbezogene Arbeitsverträge für Hochschullehrer und nicht zuletzt der verstärkte Einsatz von Lehrevaluationen zeugen von dieser Entwicklung (vgl. Hennig-Thurau/Walsh/Schrader 2004).

Die Quantifizierung von Forschungsleistungen wird sowohl für die Hochschulen als auch für das an ihnen beschäftigte wissenschaftliche Personal zunehmend wichtiger. Die Forschungsleistungen der Beschäftigten sind viel stärker als die Lehrleistungen bestimmend für die Karriere an den Universitäten und das Ansehen in der *Scientific Community*. Angesichts der großen fachlichen Spezialisierung jeder einzelnen Forscherin und jedes einzelnen Forschers<sup>1</sup> ist das Aufstellen von adäquaten und intersubjektiv nachvollziehbaren Kriterien zur Beurteilung von Forschungsleistungen naturgemäß komplex. Vor diesem Hintergrund ist es nicht verwunderlich, dass die Eignung von bibliometrischen Verfahren als quantitatives Maß für die Bewertung einer wissenschaftlichen Arbeit umstritten ist und z. T. ganz grundsätzlich kritisiert wird (vgl. Kieser 2010). Einerseits können bibliometrische Methoden als ein transparentes und intersubjektiv nachvollziehbares Maß qualitative Verfahren wie z. B. *Peer Review* ergänzen

---

\* Ich danke Alexander Dilger für wertvolle Anmerkungen und Kommentare. Für alle Inhalte und möglicherweise verbliebene Fehler ist jedoch ausschließlich der Autor selbst verantwortlich.

<sup>1</sup> Aus Gründen der besseren Lesbarkeit wird im Folgenden auf die Nennung der weiblichen Form verzichtet.

zen. Kritiker wie Lawrence (2003) sehen jedoch im Bestreben, wissenschaftliche Leistungen messbar und damit auch für Außenstehende transparenter zu machen, eine Entwicklung zur „Audit Society“ (zum Begriff siehe Power 1997).

Dieser Einwand macht deutlich, dass bibliometrischen Verfahren als Managementinstrument an Hochschulen oder in der Wissenschaftspolitik eine große Bedeutung zukommt. Ihr Einsatz birgt Chancen und Risiken: Eine gesteigerte Transparenz hinsichtlich der Leistungserbringung innerhalb der Hochschulen wie auch gegenüber ihren Stakeholdern ist geeignet, die Qualität der Managemententscheidungen zu verbessern. Für ein Berufungsverfahren ergeben sich durch den Einsatz bibliometrischer Verfahren zusätzliche Informationen, anhand derer sich die Kandidaten differenzieren lassen. Zudem entsteht ein weiterer Anreiz zu qualitativ hochwertigen Veröffentlichungen, der z. B. durch die Verknüpfung mit Leistungszulagen im Rahmen der W-Besoldung weiter gestärkt werden kann (vgl. Blomeyer 2007, Süß 2007). In dieser Wirkung sehen Kritiker aber auch die Risiken im Einsatz der Bibliometrie zu Managementzwecken. Wissenschaftler seien oftmals stark intrinsisch motiviert und bei der Implementierung von starken extrinsischen Anreizen, beispielsweise in Form von Leistungszulagen, bestehe die Gefahr der Verdrängung der intrinsischen Motivation (vgl. Frey 2003). Wenn deshalb Wissenschaftler ein stärkeres Interesse an Rangplätzen in Rankings als an der wissenschaftlichen Erkenntnis selbst entwickelten, hätten Rankings der akademischen Kultur schweren Schaden zugefügt (vgl. Weingart 2005).

## **2. Forschungsleistung und Zitationen**

### **2.1 Forschungsleistung an Hochschulen**

An den tertiären Bildungseinrichtungen im deutschsprachigen Raum nimmt die Forschung eine zentrale Rolle ein. Das Hochschulrahmengesetz (HRG) definiert sie als zentrale Aufgabe von Hochschulen: „Die Hochschulen dienen [...] der Pflege und Entwicklung der Wissenschaften und der Künste durch Forschung, Lehre, Studium und Weiterbildung“ (§ 2 HRG). Daneben wird Forschung nicht nur von Hochschulen, sondern auch von spezialisierten Forschungsinstituten (z. B. der Max-Planck-Gesellschaft) betrieben. Auch außerhalb von spezifischen Institutionen können Forschungsleistungen entstehen, wenn z. B. ein Privatgelehrter oder ein Unternehmen wissenschaftliche Erkenntnisse gewinnt.

Forschung wird im HRG als „Gewinnung wissenschaftlicher Erkenntnisse“ (§ 22 Abs. 1 HRG) definiert. Aus der Perspektive der Betriebswirtschaftslehre lässt sich Forschung an Universitäten mit Dyckhoff et al. (2005) als „Produktion neuen öffentlichen Wissens über die Welt“ bezeichnen. Diese Definition weist auf zwei wesentliche Aspekte hin, die für den hier diskutierten Kontext relevant sind: Erstens kann die Forschungstätigkeit als eine Art Produktionsprozess und damit als eine Beziehung von Input und Output betrachtet werden. Bei der Forschung werden Ressourcen (Arbeitszeit und Kapital) verbraucht und am Ende steht ein bestimmter Output, nämlich neues öffentliches Wissen. Wie dieser Output exakt aussieht und mit welchen Indikatoren er quantitativ und qualitativ gemessen werden kann, ist eine komplexe Fragestellung und Gegenstand dieses Beitrags. Zweitens sei auf den öffentlichen Charakter des Wissens hingewiesen, d. h. Wissen muss im Sinne dieser Definition nicht nur gewonnen, sondern auch publiziert werden. In dieser Hinsicht ist Forschung immer auch eine Form von Kommunikation und eine Idee, die niemanden mitgeteilt wird, noch kein vollwertiger Beitrag dazu.

Nicht öffentliche Erkenntnisse, wie sie beispielsweise in der Auftragsforschung für Industrie und Militär gewonnen werden, fallen somit nicht unter die Forschungsdefinition von Dyckhoff et al., obwohl sie mit wissenschaftlichen Methoden gewonnen sein und somit Forschungscharakter haben können. Ihnen fehlt jedoch die Öffentlichkeit, d. h. die Erkenntnisse können weder in der *Scientific Community* diskutiert werden, noch den anderen Forschern Möglichkeiten zur Erkenntnis bieten. Nicht veröffentlichte Forschungsergebnisse der Auftragsforschung stellen daher keine positive Externalität im Sinne der mikroökonomischen Theorie dar. Sie sind somit kein öffentliches, sondern ein privates Gut (vgl. grundlegend Musgrave/Musgrave 1976). Während sich der Wert der privaten Auftragsforschung unmittelbar aus der Zahlungsbereitschaft des Auftraggebers ergibt, wird die öffentliche Forschung typischerweise staatlich finanziert und die Forschenden streben nicht nur nach einer materiellen, sondern in der Regel auch nach einer ideellen Entlohnung, die sich aus der Reputation in- und außerhalb des Fachs ergibt. Auch aus der mikroökonomischen Theorie der öffentlichen Güter lässt sich also eine Differenzierung zwischen veröffentlichten und privaten Forschungsergebnissen ableiten.

## 2.2 Grundfragen der Messbarkeit

Alle Ansätze, Forschungsleistungen verschiedener Wissenschaftler miteinander zu vergleichen, sind mit spezifischen Problemen und Schwierigkeiten verbunden. Das Quantifizieren wissenschaftlicher Exzellenz mit Hilfe einer Kennzahl oder eines Rangplatzes ist in der Regel nicht unkritisch (vgl. Adler/Harzing 2009). Die Ergebnisse werden dabei oft maßgeblich von der verwendeten Methodik beeinflusst. Rein qualitative Begutachtungsansätze wie z. B. die *Peer Review* können diese Problematik aber ebenfalls nicht zufriedenstellend lösen, da sie von der Position des Gutachters abhängen und andere Gutachter typischerweise zu anderen Ergebnissen kommen. Die intersubjektive Akzeptanz solcher Ergebnisse dürfte aufgrund der im Bewertungsprozess zwangsläufig inhärenten Intransparenz tendenziell eher niedrig sein.<sup>2</sup> Nicht zuletzt sind Vergleiche auf der Grundlage einer *Peer Review* lediglich für zahlenmäßig kleine Beobachtungseinheiten praktikabel und für die Betrachtung einer größeren Zahl von Personen und Veröffentlichungen nur unter Inkaufnahme eines (ggf. prohibitiv) hohen Aufwands durchführbar. Das gilt vor allem dann, wenn man sich nicht nur auf Zeitschriftenartikel beschränkt, deren Begutachtung bereits im Einreichungsverfahren stattgefunden hat, sondern auch Monographien mit einschließen möchte.

Die quantitativen Ansätze zur Messung von Forschungsleistungen gehen davon aus, dass wissenschaftlicher Wettbewerb im Kern ein Wettbewerb um Reputation ist, in dem die Forscher danach streben, neue Erkenntnisse zu gewinnen und diese in der *Scientific Community* zu verbreiten (vgl. Klingemann 1988, sowie Dyckhoff/Schmitz 2007). Angeknüpft wird dabei ausdrücklich an der Veröffentlichung einer Erkenntnis und nicht bereits an der Erkenntnis selbst. Eine bedeutsame Entdeckung, die nicht an das (Fach-)Publikum kommuniziert wird, dürfte kaum zum wissenschaftlichen Fortschritt beitragen. Wird die Erkenntnis auf eine geeignete Art publiziert, so trägt sie in der *Scientific Community* zu weiteren neuen Entdeckungen bei, da die Forscherkollegen an die Erkenntnisse anknüpfen können.

Die Nutzung bibliometrischer Methoden zur Messung von Forschungsleistungen hat in der deutschsprachigen Betriebswirtschaftslehre noch keine lange Tradition, sondern ist ein vergleichsweise neues Konzept. Dabei konnte man sich auf Pionierarbeiten aus dem englischsprachigen Ausland und aus anderen Wissenschaften beziehen, wobei die Entwicklung von Zitationsdatenbanken und Methoden zu ihrer Auswertung in den 1960er und 70er Jahren vor-

---

<sup>2</sup> Transparent wäre das Verfahren allenfalls bei einer Veröffentlichung der Gutachten, was bei wissenschaftlichen Zeitschriften allerdings nicht üblich ist und generell im Bezug auf die Autonomie der Gutachter problematisch sein dürfte.

nehmlich von den Naturwissenschaften ausging (vgl. Osareh 1996). In diesen Disziplinen hat die Bewertung von Fachzeitschriften, Forschungsinstituten oder individuellen wissenschaftlichen Leistungen anhand bibliometrischer Indikatoren bereits eine längere Tradition, was z. T. darauf zurückzuführen ist, dass dort nahezu ausnahmslos in englischer Sprache und in internationalen Fachzeitschriften publiziert wird (vgl. Henrekson/Waldenström 2008). Auch wenn erst seit der Entwicklung von Computersystemen bibliometrische Studien in größerem Umfang möglich sind, so ist die Idee, wissenschaftliche Literatur mit statistischen Methoden zu beschreiben, deutlich älter, wie frühe Studien von Campbell (1896) und Cole/Eales (1917) belegen.

### **2.3 Zitationen als Indikator für Forschungsleistung**

Eine Grundregel des wissenschaftlichen Arbeitens besteht darin, die Nutzung fremden Gedankenguts durch einen entsprechenden Verweis darauf kenntlich zu machen. Die Autoren eines wissenschaftlichen Beitrags zitieren vornehmlich diejenigen Quellen, die ihnen einen Erkenntnisgewinn gebracht und ihre eigenen Ideen beeinflusst haben. Die Zitation spiegelt auf diese Weise die Anerkennung einer wissenschaftlichen (Vor-)Arbeit wider (vgl. Klingemann 1988, Schlinghoff/Backes-Gellner 2002). Umgekehrt deuten wenige oder völlig ausbleibende Zitationen darauf hin, dass eine Veröffentlichung aus der Sicht der *Scientific Community* keinen wesentlichen Beitrag zum Erkenntnisfortschritt geleistet hat und als eher unbedeutend angesehen werden kann (vgl. Schmitz 2008). Auf diese Weise kann die Anzahl der Zitationen als Indikator für die Qualität einer wissenschaftlichen Veröffentlichung angesehen werden (vgl. Weingart/Winterhager 1984), zumindest jedoch als Indikator für die Rezeption einer Veröffentlichung (vgl. Schmitz 2008).

Diese Sichtweise ist allerdings nicht unumstritten. MacRoberts/MacRoberts (1996, S. 442) kommen zu folgendem Urteil:

„The basic assumptions of citation analysis are clearly false. The assumption ‘that the research cited by scientists in their own papers represents a roughly valid indicator of influence on their work’ has been falsified repeatedly, as has the assumption on which it is based: that scientists are motivated to cite their influence or to give credit where credit is due.“

Dieser vernichtenden Kritik muss allerdings entgegengehalten werden, dass ein bloßes Falsifizieren der Annahmen durch die Präsentation eines geeigneten Gegenbeispiels (siehe hierzu Popper 2005) zur Klärung der Frage, ob Zitationen ein geeigneter Indikator für Forschungsleistung sind, wenig beiträgt. Entscheidend ist nicht, ob es einzelne Fälle gibt, die dem oben



Erläuterten widersprechen, sondern ob die getroffenen Annahmen die wissenschaftliche Praxis adäquat beschreiben.

Ähnliches gilt für das in der Literatur vielfach problematisierte Phänomen der negativen Zitationen. So können wissenschaftliche Arbeiten auch deswegen zitiert werden, um auf Fehler in ihnen hinzuweisen. Fraglich ist allerdings, ob dieser Umstand die Validität von Zitationsindikatoren zur Forschungsleistungsmessung tatsächlich einschränkt. Thesen, die sich als falsch erwiesen haben, werden schließlich nur dann zitiert, wenn sie trotzdem als hinreichend bedeutsam anerkannt wurden und den neuen Erkenntnissen im Weg standen. Unbedeutendes dagegen wird ignoriert (vgl. Meadows 1974). Davon abgesehen werden revolutionär neue Ansätze in der Wissenschaft typischerweise anfangs sehr scharf kritisiert, bevor sie sich durchsetzen, und ehemals wirkmächtige, aber später widerlegte Theorien haben ebenso zur Fortentwicklung des Fachs beigetragen (vgl. Garfield 1979). Letztlich ist es auch nicht Aufgabe bibliometrischer Indikatoren, den Wahrheitsgehalt einer Publikation abzubilden, sondern den Einfluss einer Veröffentlichung auf die Kommunikation des Fachs zu messen.

Empirische Studien bestätigen die hier angestellten Überlegungen zur Funktion einer wissenschaftlichen Zitation recht deutlich. So kommt Cronin (1982) bei einer Befragung im Fach Psychologie zu dem Ergebnis, dass 95 % der befragten Forscher der Aussage zustimmen, dass Anerkennung in der *Scientific Community* über das Zitieren verteilt wird. Auch andere empirische Untersuchungen weisen in dieselbe Richtung (für einen Überblick siehe Schmitz 2008). Die überwiegende Mehrheit der Wissenschaftler versteht ihre Arbeitspraktiken so, dass ein Zusammenhang von Zitation und Anerkennung einer Leistung bzw. der Qualität einer Veröffentlichung gegeben ist (vgl. Le Pair 1988). Die Tatsache, dass es in einzelnen Fällen anders sein mag, dürfte die Validität zitationsbasierter Verfahren zur Forschungsleistungsmessung nur unwesentlich relativieren.

### 3. Rankings im Hochschulwesen

#### 3.1 Rankings und die Kritik daran

Dilger (2010, S. 92) definiert Rankings als eine Methodik, bei der „Beobachtungseinheiten nach einem Kriterium in eine Reihenfolge gebracht“ werden. Diese Definition ist breit gefasst, denn es lassen sich vielerlei Dinge gemäß verschiedenen Kriterien sortieren. Beobachtungseinheit und Rangfolgenkriterium müssen also im Hinblick auf die Ziele des Rankings definiert werden. Im hier betrachteten Kontext soll die Messung von Forschungsleistungen im Mittelpunkt stehen. Beobachtungseinheiten sind daher diejenigen, die ebensolche Leistungen erbringen, namentlich die Forscher sowie im Aggregat die wirtschaftswissenschaftlichen Fachbereiche der Universitäten. Welches das zu diesem Zweck geeignete Rankingkriterium ist, soll im Folgenden diskutiert werden. Gebräuchlich sind vor allem gewichtete Zeitschriftenveröffentlichungen und Zitationskennzahlen.

Auch die Rankingziele sind umstritten, wie sich besonders in der Diskussion um die Handelsblatt-Rankings zeigte, in der neben methodologischen Anmerkungen (vgl. Hofmeister/ Ursprung 2008 für das VWL- sowie Müller 2010 für das BWL-Ranking) auch ganz grundsätzliche Kritik laut wurde (vgl. Kieser 2010). Aus der letztgenannten Position folgt ein bewusster Verzicht auf die Informationen aus quantitativen Rankingmethoden. In bestimmten Situationen müssen Rangfolgeentscheidungen jedoch zwangsläufig getroffen werden, etwa in einem Berufungsverfahren. Lässt man hier quantitative Informationen bewusst außen vor, führt dies allenfalls zu einem intransparenten und subjektiven Ranking, das methodologisch nicht mehr kritisiert werden kann, weil das Verfahren der Rangfolgenbildung nicht expliziert wird. Sachgerechter erscheint es hier, Rankings „relativ zueinander und relativ zu anderen realistischen Alternativen“ (Dilger 2010, S. 92) zu diskutieren.

Gemäß dieser Argumentation erscheint eine differenzierte Anwendung von Rankingverfahren in einer konkreten Entscheidungssituation sinnvoll. Dabei können die Ergebnisse immer auch vor dem Hintergrund der methodologischen Schwächen der jeweils verwendeten Rankingmethode diskutiert werden. Personenrankings mit vielen Beobachtungseinheiten, wie z. B. das Handelsblatt-BWL-Ranking (Müller/Storbeck 2009) oder auch die Rankings von Dilger (2010) und Dilger/Müller (2010), sollten nicht unreflektiert im Sinne einer ‚akademischen Hitparade‘ gelesen und interpretiert werden. Abgesehen von den Informationen über die namentlich genannten Bestplatzierten können nämlich auf diese Weise verschiedene Rankingmethoden exemplifiziert und kritisch diskutiert werden.

## 3.2 Methoden für Personenrankings

### 3.2.1 Zeitschriftenbasierte Personenrankings

Die zeitschriftenbasierte Methode zur Erstellung von Personenrankings ist in der BWL am meisten verbreitet und beruht auf einer Qualitätsgewichtung von Zeitschriften. Die jeweiligen Publikationen werden danach beurteilt, in welchen wissenschaftlichen Zeitschriften sie erschienen sind. Das impliziert, dass jede Zeitschrift einen gleichbleibenden Qualitätsstandard hat, der sich u. a. in der Annahme bzw. der Ablehnung von Einreichungen ausdrückt (vgl. Hennig-Thurau/Walsh/Schrader 2004). Das Zeitschriftenranking, das eine möglichst breite Abdeckung aufweisen sollte, wird auf diese Weise zur notwendigen Grundlage für ein darauf aufbauendes Personenranking.

Die zugrundeliegende Zeitschriftenliste kann auf verschiedene Bewertungsmethoden zur Ermittlung des Qualitätskriteriums zurückgreifen. Es kann z. B. auf bibliometrischen Daten oder auch auf Befragungsergebnissen basieren. Bibliometrische Zeitschriftenrankings beruhen in der Regel auf *Impact*-Faktoren, mit denen die durchschnittliche Zitationshäufigkeit der in den jeweiligen Zeitschriften erschienenen Beiträge abgebildet wird. Ein Beispiel hierfür ist die Zeitschriftenliste des *Social Sciences Citation Index* (SSCI). Das VHB-Jourqual ist in Deutschland das bekannteste befragungsbasierte Zeitschriftenranking, für das die Mitglieder des Verbands der Hochschullehrer für Betriebswirtschaft jeweils die Zeitschriften ihrer Teildisziplin einschätzen sollten (vgl. Hennig-Thurau/Walsh/Schrader 2004).

Ein entscheidender Vorteil zeitschriftenbasierter Personenrankings liegt in der vergleichsweise unproblematischen Erstellung, wenn bereits ein Zeitschriftenranking vorliegt. Ausgehend von den Qualitätskennzahlen der jeweiligen Zeitschriften können die Publikationen der Wissenschaftler mit Gewichten versehen und aufaddiert werden. Dieses Verfahren ist transparent und für den einzelnen Wissenschaftler ex ante gut abschätzbar. Aus der Qualitätsgewichtung der Zeitschriften, die als Qualitätsindikator auf die einzelnen Veröffentlichungen darin übertragen wird, ergibt sich ein unmittelbarer Anreiz, in hochrangigen Zeitschriften zu publizieren.

Ein diskretionärer Spielraum besteht dann, wenn mehrere Zeitschriftenrankings miteinander kombiniert werden, die auf jeweils unterschiedlichen Grundlagen basieren und eine unterschiedlich breite Abdeckung aufweisen. Auf diese Problematik ist bereits an anderer Stelle in Bezug auf das Handelsblatt-BWL-Ranking hingewiesen worden (vgl. Müller 2010).

Zeitschriftenbasierte Personenrankings sind allerdings konstruktionsbedingt mit einer Reihe von Problemen behaftet. Erstens können nur in Zeitschriften publizierte Forschungsbeiträge für das Ranking verwendet werden. Monographien, Beiträge in Sammelbänden oder Online-Publikationen bleiben methodenbedingt unberücksichtigt. Weil auch in diesen Medien für die Fortentwicklung der BWL bedeutende Forschungsergebnisse veröffentlicht werden (vgl. Griffith/Cavusgil/Xu 2008, Neuhaus/Daniel 2008), ist die exklusive Betrachtung von Zeitschriftenbeiträgen inhaltlich nicht zu rechtfertigen. Sie wird in der Regel damit verteidigt, dass für andere Publikationstypen vergleichbare Qualitätsindikatoren fehlen (vgl. Hennig-Thurau/Walsh/Schrader 2004, S. 521). Diese Argumentation verliert allerdings dann an Gewicht, wenn die Messung von Qualitätsindikatoren (z. B. durch Zitationen) auf alle Publikationsarten ausgedehnt werden kann. Dies ist heute mit umfassenden Datenbanken wie z. B. *Google Scholar* durchaus möglich.

Zweitens bedeutet die Bewertung eines Forschungsbeitrags danach, in welchem Medium er erschienen ist, immer die Zuweisung eines Mittelwerts. Der Wert der Zeitschrift wird durch die Zitationen bestimmt, die auf die in ihr publizierten Beiträge entfallen. Der so im Zeitschriftenranking gemittelte *Impact*-Faktor wird nun wieder den einzelnen Beiträgen zugeordnet. Auf diese Weise werden die verfügbaren Informationen nicht bestmöglich ausgenutzt. Zeitschriftenbasierte Personenrankings enthalten eine unnötige Unschärfe an dieser Stelle. Wirklich herausragende Publikationen werden auf diese Weise tendenziell unterschätzt. Seglen (1994) hat nachgewiesen, dass der *Impact*-Faktor einer Zeitschrift nur sehr schwach mit dem *Impact* der darin enthaltenen Beiträge korreliert ist. Die Zitationen der Artikel bilden das Renommee der Zeitschrift, aber nicht umgekehrt (siehe auch Seglen 1997). So sind selbst in erstrangigen Zeitschriften viele Beiträge zu finden, die nur sehr wenig oder überhaupt nicht rezipiert wurden, während für die weitere Forschung bedeutende Erkenntnisse mitunter auch in zweitrangigen Journals publiziert wurden (vgl. Frey/Rost 2010).

Im Fall von Personenrankings, die auf bibliometrisch erzeugten Zeitschriftenrankings basieren, stellen sich darüber hinaus weitere Probleme. Weisen die Zitationsdatenbanken größere systematische Lücken auf, können daraus Verzerrungen auch auf der Ebene des Personenrankings resultieren. Deutlich wird das, wenn man sich z. B. die *Impact*-Faktoren der Zeitschriften im SSCI ansieht. Der SSCI enthält kaum deutschsprachige Zeitschriften der BWL, die *Betriebswirtschaftliche Forschung und Praxis* (BFuP) ist eine der wenigen Ausnahmen. Clermont/Schmitz (2008, S. 23) folgern daraus, dass es mit dem SSCI „nicht möglich [sei], deutschsprachig publizierte Forschungsbeiträge oder ihre Zitationen [...] sinnvoll zu erheben“.

Weil wissenschaftliche Diskurse darüber hinaus oft regionale Schwerpunkte haben, werden auch die *Impact*-Faktoren der wenigen im SSCI enthaltenen deutschsprachigen Zeitschriften niedriger ausgewiesen, als sie in einer vollständigen Datenbank wären. Da deutschsprachige Beiträge vornehmlich von deutschsprachigen Beiträgen zitiert werden, entsteht eine systematische Verzerrung zulasten der deutschsprachigen Literatur. So ‚fehlen‘ den Beiträgen in der BFuP beispielsweise die Zitationen aus der nicht im SSCI enthaltenen *Zeitschrift für Betriebswirtschaft* (ZfB). Auf diese Weise erscheinen auf Deutsch publizierte Beiträge weniger wirkmächtig als englischsprachige.

Die Kombination mehrerer unvollständiger Zeitschriftenrankings zu einer gemeinsamen Liste (für einen methodologisch anspruchsvollen Ansatz vgl. Schulze/Warning/Wiermann 2008) zwecks Erstellung eines Personenrankings löst das Problem nur unvollständig (vgl. Albers 2009 zur Kritik). Zwar entsteht auf diese Weise die gewünschte Abdeckungsbreite. Die systematische Verzerrung zu Ungunsten der nicht-englischsprachigen Literatur verschwindet jedoch nicht. Werden die deutschen Titel in jedem Ranking systematisch unterhalb ihres eigentlichen (d. h. bei einer vollständigen Datenbank sich ergebenden) Werts aufgeführt, so werden sie auch in der aggregierten Zeitschriftenliste nicht auf adäquaten Plätzen liegen.

Darüber hinaus können sich die dem Personenranking zugrunde liegenden Zeitschriftenrankings im Laufe der Zeit verändern. *Impact*-Faktoren sind typischerweise nicht konstant, sondern zeichnen die Entwicklung einer Zeitschrift nach. Da auch vor vielen Jahren erschienene Beiträge jeweils mit den aktuellen *Impact*-Faktoren (oder Befragungsdaten) der Zeitschriften bewertet werden, entstehen im Laufe der Zeit zufällige Effekte zu Gunsten oder zu Ungunsten der damaligen Autoren. Diese Auf- oder Abwertungen haben in Bezug auf die zu bewertenden alten Publikationen keine reale Substanz: Wird beispielsweise eine Zeitschrift von VHB-Jourqual 1 zu Jourqual 2 heruntergestuft, so sagt dies wenig über einen vor zehn Jahren in ihr erschienenen Beitrag aus.

Schließlich ist auch beim zeitschriftenbasierten Personenranking im schlimmsten Fall mit strategischem Zitieren zu rechnen (vgl. Dilger 2000). Die Autoren einer Zeitschrift haben ein Interesse daran, ebendiese in der Folgezeit häufig zu zitieren, um über den *Impact*-Faktor der Zeitschrift die Bewertung ihrer eigenen Beiträge zu verbessern. Aufgrund von Kollektivgutproblemen ist der Anreiz zur Selbstzitation bei zeitschriftenbasierten Personenrankings allerdings geringer als bei zitationsbasierten. Zeitschriftenrankings auf der Basis von Expertenurteilen reizen ein ähnliches Verhalten an: Wissenschaftler sind möglicherweise dazu verleitet, die Zeitschriften mit ihren eigenen Beiträgen besonders gut zu bewerten (vgl. Breuer 2009).

Aus diesem Grund kommt der Anzahl der für jede Zeitschrift abgegebenen Einschätzungen z. B. beim Jourqual 2 eine wichtige Bedeutung zu.

### 3.2.2 Zitationsbasierte Personenrankings

Wird die Zitation als Indikation für die Rezeption und Anerkennung einer Forschungsleistung verstanden, so kann anhand dieser Kennzahl ein Ranking konstruiert werden. Nicht der Umstand der Veröffentlichung bzw. die Platzierung eines Beitrags in einem bestimmten wissenschaftlichen Journal wird als Qualitätsgewicht herangezogen, sondern es wird direkt auf den *Impact* des einzelnen Beitrags abgestellt. Die Anzahl der Zitationen als Qualitätsindikator zu nutzen ist kongruent mit dem oben erläuterten Konzept des wissenschaftlichen Wettbewerbs. Qualität wird nicht durch das Urteil des Herausgebers einer Zeitschrift (und der ggf. von ihm bestellten Gutachter), sondern durch die Reaktion der gesamten *Scientific Community* abgebildet.

Weil die Zitationen direkt für jede einzelne Veröffentlichung jedes Forschers erhoben werden, entfällt die im Falle der zeitschriftenbasierten Rankings kritisierte Mittelwertbildung. Es wird nur der individuelle *Impact* jedes einzelnen Beitrags gemessen. Externe Effekte von anderen im gleichen Medium veröffentlichten Beiträgen existieren höchstens indirekt, wenn eine hochrangige Zeitschrift generell mehr Leser hat und ein Beitrag darin ein größeres Publikum und mehr potenzielle Rezipienten erreicht.

Ein zweiter wichtiger Vorteil, den zitationsbasierte Rankings gegenüber zeitschriftenbasierten haben, liegt im grundsätzlich breiteren Fokus der Betrachtung. Die Rankingmethode ist nicht auf Zeitschriftenbeiträge begrenzt, sondern kann auch Monographien und Beiträge in Sammelbänden umfassen, sofern die Zitationen dieser Veröffentlichungen erhoben werden können. Die Qualität des zitationsbasierten Rankings ist daher eng mit der Qualität und der Abdeckungsbreite der zugrundeliegenden Datenbank verknüpft. Der Marktführer im Bereich der wissenschaftlichen Datenbanken, das *ISI Web of Science* (SSCI, SCI-X und A&HCI), erfüllt diese Anforderungen nur unvollständig. Es wertet alle Zitationen aus den im Index enthaltenen Zeitschriften aus, d. h. auch solche, die auf Buchpublikationen verweisen. Monographien und Sammelbände werden erfasst, sofern sie in den jeweiligen Journals zitiert wurden. Ein Problem liegt aber in der mangelhaften Abdeckung im Bereich der deutschsprachigen Zeitschriften, was zu systematischen Verzerrungen führt (vgl. Clermont/Schmitz 2008). Dilger

(2000) forderte daher auch aus diesem Grund die Entwicklung eines deutschsprachigen Zitationsindex, der aber bis heute nicht existent oder konkret in Planung ist.

In Ermangelung eines speziell für die Zwecke der deutschsprachigen Sozialwissenschaften zugeschnittenen Zitationsindex kann *Google Scholar* als am ehesten geeignete Alternative gelten. Die Datenbank ist leistungsstark und kostenlos und wurde schon für eine Reihe von Studien erfolgreich eingesetzt (vgl. Henrekson/Waldenström 2008, Dilger/Müller 2010, Müller 2010, Breuer 2009). Aufgrund der breiten Abdeckung wissenschaftlicher Publikationen werden Zitationen von allen in alle Medientypen abgebildet. Auch Buch-zu-Buch-Verweise werden erfasst, die beim *ISI Web of Science* konstruktionsbedingt unberücksichtigt bleiben. Weil Veröffentlichungsdatenbanken von Wissenschaftsverlagen und Daten aus dem Internet automatisch durchsucht und ausgewertet werden, enthält *Google Scholar* in Bezug auf die Betriebswirtschaftslehre mehr Publikationen als die direkten Konkurrenten *ISI Web of Science* oder *Scopus* (vgl. Harzing/van der Wal 2008). Das Abfragen der Informationen kann mittels des von Anne-Wil Harzing entwickelten und kostenlos verfügbaren Programms *Publish or Perish* vorgenommen werden.<sup>3</sup> Es bietet eine benutzerfreundliche Bedienoberfläche und errechnet zahlreiche bibliometrische Kennzahlen.

The screenshot shows the 'Harzing's Publish or Perish' software interface. The main window is titled 'Multi-query center' and contains a table of queries with columns for Name, Papers, Cites, Cites..., Auth..., h, g, and Query Date. Below this is a 'Results' section with summary statistics for Papers, Citations, Years, Cites/year, Cites/paper, Cites/author, Papers/author, Authors/paper, h-index, g-index, hc-index, hi-index, AWCR, AWCRpA, e-index, and hm-index. At the bottom, there is a table of results with columns for Cites, Per year, Rank, Authors, Title, Year, and Publication. The status bar at the bottom indicates the version '2.8.3644' and the date 'Donnerstag, 21. Januar 2010'.

**Abb. 1: Zitationsanalyse mit Publish or Perish**

<sup>3</sup> Kostenloser Download unter [www.harzing.com/pop.html](http://www.harzing.com/pop.html) (letzter Abruf am 25.01.2012).

Gegen zitationsbasierte Rankings werden in der Literatur ebenfalls einige Punkte vorgebracht: Ein im Rahmen der Zitationsanalyse häufig zu beobachtendes Phänomen ist der Matthäus-Effekt (vgl. Merton 1968).<sup>4</sup> Er beschreibt die Tatsache, dass bereits erfolgreiche Forscher besonders häufig zitiert werden, da sich die Zitierenden gern auf eine Autorität des Fachs beziehen. Daneben scheint es im Zitierverhalten eine Präferenz für methodologische Publikationen und Überblickswerke zu geben (vgl. Voeth/Gawantka/Chatzopoulou 2006). Beide Einwände richten sich bei näherer Betrachtung aber nicht gegen das zitationsbasierte Ranking im Speziellen, sondern berühren eine grundlegende Frage der Mess- und Vergleichbarkeit wissenschaftlicher Leistungen. Auch bei zeitschriftenbasierten Rankings würden diese Phänomene nämlich bereits etablierte und methodologisch ausgerichtete oder Überblicksartikel enthaltende Zeitschriften begünstigen. Das gilt insbesondere dann, wenn die Zeitschriftenbewertung auf *Impact*-Faktoren basiert. Ähnliches kann aber auch für befragungsbasierte Zeitschriftenrankings plausibel angenommen werden.

Da Rankings naturgemäß vergangenheitsbezogen sind, gibt es eine Präferenz für ältere Veröffentlichungen, weil diese bereits auf eine längere Zitationsgeschichte zurückblicken. Das ist bei der Wahl eines geeigneten Untersuchungszeitraums zu berücksichtigen, sollte jedoch nicht voreilig als Fehler oder Verzerrung der Methodik interpretiert werden. Es ist vielmehr ein grundlegendes Charakteristikum von Wissenschaft, dass Ideen und Erkenntnisse erst im Nachhinein in ihrer wahren Tragweite abgeschätzt werden können.<sup>5</sup> Eine wissenschaftliche Veröffentlichung geht vielmehr durch einen Lebenszyklus, zu dessen Beginn die Zitationen stark zunehmen, einen Höhepunkt erreichen und dann langsam wieder fallen, weil die veröffentlichten Inhalte entweder keine Relevanz mehr haben oder in das Allgemeinwissen des Fachs übergegangen sind und deswegen nicht mehr zitiert werden müssen (vgl. Aizenman/Kletzer 2008).

Ein praktischer Nachteil zitationsbasierter Rankings ist der im Vergleich zu zeitschriftenbasierten Rankings ungleich größere Aufwand, den ihre Erstellung erfordert (vgl. Dilger 2010): Erstens müssen für jede Publikation Zitationskennzahlen aus der Datenbank entnommen werden. Zweitens müssen sämtliche Informationen für jede Neuauflage des Rankings komplett neu erhoben werden, da auch alle älteren Veröffentlichungen weiter zitiert werden können. Automatisierte Software-Tools sind gegenwärtig weder für das *ISI Web of Science* noch für *Google Scholar* verfügbar.

---

<sup>4</sup> Benannt nach Mt. 25, 29: „Denn wer da hat, dem wird gegeben werden, dass er in Fülle habe“.

<sup>5</sup> So wird auch der sogenannte Wirtschaftsnobelpreis typischerweise für Erkenntnisse verliehen, die bereits seit Jahrzehnten publiziert sind.



Ein gravierender Nachteil zitationsbasierter Rankings liegt in ihrer Anreizstruktur im Hinblick auf mögliches strategisches Verhalten. Selbstzitationen und Zitierkartelle könnten von Wissenschaftlern dazu genutzt werden, die eigene Rankingposition zu verbessern (vgl. Dilger 2000). Die Datenbanken können zwar Selbstzitationen grundsätzlich filtern. In *Google Scholar* geschieht dies aber – trotz anderweitiger Behauptungen von *Google* –, wenn überhaupt, nur sehr unzureichend (vgl. Henrekson/Waldenström 2008). Zitierkartelle sind hingegen technisch nicht zu erkennen und im Übrigen von korrektem wissenschaftlichen Verhalten auch kaum trennscharf abzugrenzen: Wenn sich nur sehr wenige Forscher mit einem eng umgrenzten Thema beschäftigen und sich in ihren Publikationen aufeinander beziehen, deutet dies nicht zwangsläufig auf ein Zitierkartell hin (vgl. Dyckhoff/Thieme/Schmitz 2005). Zum gegenwärtigen Zeitpunkt dürften die Anreizwirkungen zitationsbasierter Rankings ohnehin noch keine große Rolle spielen, da sie in den deutschsprachigen Sozialwissenschaften sehr ungebrauchlich sind. Erwartet werden muss strategisches Verhalten erst dann, wenn ein zitationsbasiertes Ranking große Aufmerksamkeit findet und das Abschneiden darin Konsequenzen für die Gerankten hat.

### **3.2.3 Weitere Methoden**

Neben den beiden hier vorgestellten Verfahren sind weitere Kriterien denkbar, nach denen Personen oder Hochschulen in eine Reihenfolge gebracht werden können. Für ein Fachbereichsranking in der BWL ließe sich z. B. auch der befragungsbasierte Ansatz analog zum VHB-Jourqual Zeitschriftenranking anwenden.

### ***CHE-Forschungsranking als multidimensionales Verfahren***

Im Rahmen des multidimensionalen Forschungsrankings des Centrums für Hochschulentwicklung (CHE) wurde die Befragungsmethode als eine der mehreren Dimensionen gewählt (vgl. hierzu im Folgenden Berghoff et al. 2009). BWL-Professoren wurden dazu befragt, wie sie die Reputation der anderen Hochschulen in der Forschung einschätzen. Als weitere Indikatoren für Forschungsleistung wurde u. a. die Anzahl an Promotionen sowie die Höhe und Struktur der verausgabten Drittmittel herangezogen. In Bezug auf die Promotionen kann hinterfragt werden, ob die Betreuung von Doktoranden nicht eher eine Form von Lehrleistung als ein Forschungsindikator ist. Zudem wird die Qualität der Promotionen nicht berücksichtigt.

Hinsichtlich der Drittmittel ist unklar, ob es sich überhaupt um ein Leistungsmaß handelt. Versteht man wissenschaftliche Forschung betriebswirtschaftlich als Produktionsprozess von Wissen, so sind sie als Inputgröße anzusehen und sollten bei einer Outputmessung unberücksichtigt bleiben (vgl. Rassenhövel/Dyckhoff 2006).

Weitere Indikatoren des CHE-Rankings waren in der Betriebs- und Volkswirtschaftslehre deutschsprachige und internationale Publikationen. Im Rahmen der Publikationsanalyse wurde ein gemischtes Verfahren aus zeitschriftenbasierter Qualitätsgewichtung und reiner Zählung der Beiträge angewandt. Vor dem Hintergrund der eingangs erläuterten Überlegungen kann ein solches Vorgehen durchaus hinterfragt werden. So ist z. B. nicht einsichtig, warum internationale und nationale Publikationen zwei getrennte Indikatoren darstellen sollen. Inhaltlich unklar bleibt ferner, wieso erstere lediglich gezählt, letztere aber auf Koautorenschaften und Seitenzahl normiert und anhand des VHB-Jourqual qualitätsgewichtet werden.

Alle Indikatoren des CHE-Rankings werden schließlich in einem *Scoring*-Modell aggregiert, das die Forschungsleistung der betriebswirtschaftlichen Fachbereiche multidimensional abbilden soll. Dabei fällt auf, dass die Korrelationen zwischen den unterschiedlichen Indikatoren vielfach nicht signifikant und auch sonst meist nur sehr schwach ausgeprägt sind (vgl. Berghoff et al. 2009). Vor dem Hintergrund dieser Ergebnisse scheint ein gewichtetes Aufaddieren der in verschiedene Richtungen deutenden Indikatoren nur bedingt sinnvoll: Aufgrund der Divergenz der Zwischenergebnisse ist zu vermuten, dass das resultierende Ranking stark von der Gewichtung im *Scoring*-Modell abhängt. In jedem Fall sind die Ergebnisse aufgrund der Vielzahl der verwendeten und miteinander verknüpften Methoden kaum noch transparent zu machen. Das hier nur in seinen grundsätzlichen Zügen wiedergegebene Verfahren verdeutlicht, dass eine Vielzahl möglicher Indikatoren zur Messung von Forschungsleistung existiert. Daneben sind noch viele weitere Kriterien vorstellbar. Vor diesem Hintergrund scheint eine begründete Auswahl sinnvoller als der Versuch, sämtliche möglichen Verfahren in ein Ranking zu implementieren.

Die anhand des CHE-Rankings vorgestellten Alternativindikatoren zur Messung von Forschungsleistung ließen sich grundsätzlich nicht nur auf Hochschulen, sondern auch auf Lehrstühle, Institute oder auf Personen anwenden. Anhand von Drittmittelverwendung, Anzahl der Promotionen, mittels einer Befragung (oder analog zum CHE-Modell durch eine Kombination aller genannten Verfahren) ließen sich Personenrankings erstellen. Ob eine solches Vorgehen zielführend ist, kann mit Blick auf die Ausführungen zum CHE-Hochschulranking allerdings stark bezweifelt werden.

### *Herausgebertätigkeit als Qualitätsindikator*

Einen anderen Ansatz wählen Frey/Rost (2010), indem sie wissenschaftliche Qualität durch die Herausgebertätigkeit eines Forschers für hochrangige wissenschaftliche Zeitschriften (die im ISI *Web of Knowledge Journal Citation Report* oder im VHB-Jourqual Ranking mit AA+ bis B bewertet wurden) abbilden. Wenngleich man den Autoren darin zustimmen muss, dass auch die Administration und Redaktion von wissenschaftlichen Zeitschriften eine wichtige Aufgabe eines Hochschullehrers ist, so bleibt letztlich offen, wieso diese Methodik den an den Publikationen der Forscher orientierten Verfahren überlegen sein soll. Betrachtet man die Produktion von Wissen als primären Prozess einer Hochschule (neben der Lehre) oder einer wissenschaftlichen Forschungseinrichtung, so käme der Organisation des Publikationsverfahrens eher die Rolle eines Sekundär- oder Hilfsprozesses zu (vgl. Porter 1985). Zwar soll nicht bestritten werden, dass sich die qualifiziertesten Forscher eines Fachs besonders für die Tätigkeit als Herausgeber einer Zeitschrift empfehlen. Dennoch sollte sich die Messung von Forschungsleistung so eng wie möglich am Inhalt bzw. den wissenschaftlichen Erkenntnissen an sich und weniger an den dazu notwendigen Hilfsprozessen orientieren.

Von diesen eher konzeptionellen Überlegungen abgesehen hat das von Frey/Rost entworfene Verfahren praktisch wenig Aussagekraft. So gibt es in einem weltweiten Ranking der führenden Ökonomen nur sieben (sic!) mögliche Rangplätze. Angewandt auf die deutschsprachige Betriebswirtschaftslehre ließe sich vermuten, dass sich nahezu alle Forscher auf einem geteilten siebten Platz wiederfänden. Der Ansatz scheint also in Bezug auf die Messung von Forschungsleistungen einer Vielzahl von Forschern ungeeignet. Eine mögliche Erweiterung der untersuchten Herausgeberkreise könnte zwar für mehr Forscher zu Ergebnissen führen, wirft aber die Frage auf, wie die Herausgeberschaften in unterschiedlichen Zeitschriften gegeneinander gewichtet werden sollen. So wäre die Tätigkeit als Herausgeber von beispielsweise *WISU* sicherlich weniger reputierlich als die für das *Journal of Economic Literature*. Diese Frage wird von Frey/Rost nicht gelöst, sondern durch die Begrenzung auf wenige internationale Top-Journals umgangen.

## 4. Fazit

Die vergleichende Analyse der methodologischen Ansätze zur Erstellung von Forschungsleistungsrankings hat gezeigt, dass die Betrachtung der Zitationen einer Veröffentlichung gegenüber den anderen besprochenen Methoden zwei wesentliche konzeptionelle Vorteile aufweist: Erstens ist dieses Verfahren nicht notwendigerweise auf Zeitschriftenpublikationen begrenzt, sondern es können auch Monographien und Sammelbände betrachtet werden, und zweitens können die verfügbaren Daten bestmöglich ausgenutzt werden, weil direkt an der einzelnen Publikation angeknüpft wird. Diese konzeptionellen Vorteile dürften die praktischen Nachteile, die in den Anforderungen an die Datenbank und im Erstellungsaufwand liegen, langfristig mehr als aufwiegen.

Die Eignung einer Methode muss jedoch immer vor dem Hintergrund der Ziele des jeweils aufzustellenden Rankings beurteilt werden. Soll beispielsweise ein Bewerberranking als Hilfsmittel zur (Vor-)Auswahl im Rahmen eines Berufungsverfahrens erstellt werden, so erscheint der zitationsbasierte Ansatz als der am ehesten geeignete. Der im Vergleich zum zeitschriftenbasierten Ranking höhere Aufwand dürfte bei einer begrenzten Zahl von Personen nicht ins Gewicht fallen und auch die Eignung der dazu heranzuziehenden Zitationsdatenbanken lässt sich anhand der Abdeckungsbreite leicht beurteilen, wenn man die Suchergebnisse mit den Publikationslisten der Bewerber abgleicht. Fehler in der Datenbank, wie sie insbesondere für *Google Scholar* dokumentiert sind (vgl. Jacsó 2006), lassen sich unmittelbar erkennen und oft bereits mit der Auswertungssoftware *Publish or Perish* manuell korrigieren. Dennoch mag der zitationsbasierte Ansatz nicht in jedem Fall die Methode der Wahl sein. Die auf Zeitschriftengewichten basierende Methode zeichnet sich nämlich durch eine ex ante größere Planbarkeit aus, wie sie z. B. bei Vereinbarungen über Leistungszulagen im Rahmen der W-Besoldung gegeben sein sollte. Wenn eine Hochschulleitung derartige Instrumente trotz der mit ihnen verbundenen Probleme in der Professorenbesoldung nutzen möchte, müssen die vereinbarten Ziele von den Forschern sinnvoll angestrebt werden können, was bei einem Artikel in einer bestimmten Zeitschriftenkategorie im Gegensatz zu einer definierten Zahl von Zitationen eher möglich erscheint.

Da die anzuwendenden Methoden vor dem Hintergrund unterschiedlicher Ziele bewertet werden sollen, folgt daraus ein Nebeneinander verschiedener Rankings. Keine Methodik dominiert bislang sämtliche Alternativen in allen Eigenschaften und folglich dürfen die Ergebnisse immer nur innerhalb der methodeninduzierten Grenzen interpretiert werden. Wenn metho-

disch unterschiedliche Rankings mit voneinander abweichenden Ergebnissen nebeneinander existieren, kann dies jedoch für den Fortschritt der Bibliometrie durchaus hilfreich sein, denn eine weitere Diskussion und die Weiterentwicklung der Methoden sollten sich auf die Validität der Messinstrumente positiv auswirken. Nicht zuletzt ist ein Nebeneinander mehrerer gleichberechtigter Ansätze auch dazu geeignet, die von Kritikern befürchteten Fehlanreize in gewissen Grenzen zu halten. Die befürchtete Zielverschiebung der Gerankten weg von wissenschaftlicher Exzellenz hin zu einer Optimierungsstrategie für ein bestimmtes Ranking ist aus der Sicht der Betroffenen weniger attraktiv, wenn es mehrere Rankings mit verschiedenen Metriken gibt. In diesem Sinne ist die Entwicklung alternativer und innovativer bibliometrischer Methoden weiterhin sehr begrüßenswert.

## Literatur

- Adler, N.J./Harzing, A. (2009): „When knowledge wins: Transcending the sense and non-sense of academic rankings“, *The Academy of Management Learning and Education* 8, S. 72-95.
- Aizenman, J./Kletzer, K. (2008): „The Life Cycle of Scholars and Papers in Economics: the 'Citation Death Tax'“, NBER Working Paper No. 13891, Cambridge (MA).
- Albers, S. (2009): „Misleading rankings of research in business“, *German Economic Review* 10, S. 352-363.
- Berghoff, S./Federkeil, G./Giebisch, P./Hachmeister, C./Hennings, M./Roessler, I./Ziegele, F. (2009): „Das CHE-Forschungsranking deutscher Universitäten 2009“, Gütersloh: Centrum für Hochschulentwicklung.
- Blomeyer, C. (2007): „Professorenbesoldung“, Hamburg: Dashöfer.
- Breuer, W. (2009): „Google Scholar as a Means for Quantitative Evaluation of German Research Output in Business Administration: Some Preliminary Results“, Diskussionspapier, Aachen, im Internet unter <http://ssrn.com/abstract=1280033> (letzter Aufruf am 25.11.2011).
- Campbell, F. (1896): „The Theory of National and International Bibliography“, London: Library Bureau.
- Clermont, M./Schmitz, C. (2008): „Erfassung betriebswirtschaftlich relevanter Zeitschriften in den ISI-Datenbanken sowie der Scopus-Datenbank“, *Zeitschrift für Betriebswirtschaft* 78, S. 987-1010.
- Cole, F.J./Eales, N.B. (1917): „A History of Comparative Anatomy: Pt. 1. A Statistical Analysis of the Literature“, *Science Progress* 11, S. 578-596.

- Cronin, B. (1982): „Norms and functions in citation: The view of journal editors and referees in psychology“, *Social Science Information Studies* 2, S. 65-77.
- Dilger, A. (2010): „Rankings von Zeitschriften und Personen in der BWL“, *Zeitschrift für Management* 5, S. 91-102.
- Dilger, A. (2000): „Plädoyer für einen Sozialwissenschaftlichen Zitationsindex“, *Die Betriebswirtschaft* 60, S. 473-484.
- Dilger, A./Müller, H. (2010): „A citation based ranking of German-speaking researchers in business administration with data of Google Scholar“, *Diskussionspapier des Instituts für Organisationsökonomik* 1, Münster.
- Dyckhoff, H./Rassenhövel, S./Gilles, R./Schmitz, C. (2005): „Beurteilung der Forschungsleistung und das CHE-Forschungsranking betriebswirtschaftlicher Fachbereiche“, *Wirtschaftswissenschaftliches Studium* 34, S. 62-69.
- Dyckhoff, H./Schmitz, C. (2007): „Forschungsleistungsmessung mittels SSCI oder SCI-X? Internationale Sichtbarkeit und Wahrnehmung der Betriebswirtschaftslehre von 1990 bis 2004“, *Die Betriebswirtschaft* 67, S. 638-662.
- Dyckhoff, H./Thieme, A./Schmitz, C. (2005): „Die Wahrnehmung deutschsprachiger Hochschullehrer für Betriebswirtschaft in der internationalen Forschung: Eine Pilotstudie zu Zitationsverhalten und möglichen Einflussfaktoren“, *Die Betriebswirtschaft* 65, S. 350-372.
- Frey, B.S. (2003): „Publishing as prostitution? Choosing between one's own ideas and academic success“, *Public Choice* 116, S. 205-223.
- Frey, B.S./Rost, K. (2010): „Do rankings reflect research quality?“, *Journal of Applied Economics* 13, S. 1-38.
- Garfield, E. (1979): „Is citation analysis a legitimate evaluation tool?“, *Scientometrics* 1, S. 359-375.
- Griffith, D.A./Cavusgil, S.T./Xu, S. (2008): „Emerging themes in international business research“, *Journal of International Business Studies* 39, S. 1220-1235.
- Harzing, A./van der Wal, R. (2008): „Google Scholar as a new source for citation analysis“, *Ethics in Science and Environmental Politics* 8, S. 61-73.
- Hennig-Thurau, T./Walsh, G./Schrader, U. (2004): „VHB-JOURQUAL: Ein Ranking von betriebswirtschaftlich-relevanten Zeitschriften auf der Grundlage von Expertenurteilen“, *Zeitschrift für betriebswirtschaftliche Forschung* 56, S. 520-545.
- Henrekson, M./Waldenström, D. (2008): „How Should Research Performance be Measured? Evidence from Rankings of Academic Economists“, *Working Paper Series in Economics and Finance*, Stockholm.
- Hofmeister, R./Ursprung, H.W. (2008): „Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung“, *Perspektiven der Wirtschaftspolitik* 9, S. 254-266.
- Jacsó, P. (2006): „Deflated, inflated and phantom citation counts“, *Online information review* 30, S. 297-309.

- Kieser, A. (2010): „Unternehmen Wissenschaft?“, *Leviathan* 38, S. 347-367.
- Klingemann, H. (1988): „Zitierhäufigkeit als Qualitätsindikator: Eine Rangordnung der amerikanischen politikwissenschaftlichen Fachbereiche in den 80er Jahren“, in: Daniel, H.D./Fisch, R. (Hg.): „Evaluation von Forschung“, Konstanz: Universitätsverlag, S. 201-214.
- Lawrence, P.A. (2003): „The politics of publication“, *Nature* 422, S. 259-261.
- Le Pair, C. (1988): „The citation gap of applicable science“, in: van Raan, A.F.J. (Hg.): „Handbook of Quantitative Studies of Science and Technology“, Amsterdam: Elsevier/North-Holland, S. 537-553.
- MacRoberts, M.H./MacRoberts, B.R. (1996): „Problems of citation analysis“, *Scientometrics* 36, S. 435-444.
- Meadows, A.J. (1974): „Communication in science“, London: Butterworths.
- Merton, R.K. (1968): „The Matthew effect in science“, *Science* 159, S. 56-63.
- Müller, A./Storbeck, O. (2009): „Ranking: Die BWL erfindet sich neu“, *Handelsblatt* vom 20.5.
- Müller, H. (2010): „Wie valide ist das Handelsblatt-BWL-Ranking: Zeitschriften und zitationsbasierte Personenrankings im Vergleich“, *Betriebswirtschaftliche Forschung und Praxis* 62, S. 152-166.
- Neuhaus, C./Daniel, H. (2008): „Data sources for performing citation analysis: an overview“, *Journal of Documentation* 64, S. 193-210.
- Osareh, F. (1996): „Bibliometrics, citation analysis and co-citation analysis: A review of literature I“, *Libri* 46, S. 149-158.
- Popper, K. (2005): „Logik der Forschung“, 11. Auflage, Tübingen: Mohr Siebeck.
- Porter, M.E. (1985): „Competitive advantage: creating and sustaining superior performance“, New York: Free Press.
- Power, M. (1997): „The audit society: rituals of verification“, New York: Oxford University Press.
- Rassenhövel, S./Dyckhoff, H. (2006): „Die Relevanz von Drittmittelindikatoren bei der Beurteilung der Forschungsleistung im Hochschulbereich“, in: Zelewski, S./Akca, N. (Hg.): „Fortschritt in den Wirtschaftswissenschaften: Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen“, Wiesbaden: Deutscher Universitätsverlag, S. 85-112.
- Schlinghoff, A./Backes-Gellner, U. (2002): „Publikationsindikatoren und die Stabilität wirtschaftswissenschaftlicher Zeitschriftenrankings“, *Zeitschrift für betriebswirtschaftliche Forschung* 54, S. 343-362.
- Schmitz, C. (2008): „Messung der Forschungsleistung in der Betriebswirtschaftslehre auf Basis der ISI-Zitationsindizes: Eine kritische Analyse anhand konzeptioneller Überlegungen und empirischer Befunde“, Lohmar/Köln: Eul Verlag.

- Schulze, G.G./Warning, S./Wiermann, C. (2008): „Zeitschriftenrankings für die Wirtschaftswissenschaften? Konstruktion eines umfassenden Metaindexes“, *Perspektiven der Wirtschaftspolitik* 9, S. 286-305.
- Seglen, P.O. (1997): „Why the impact factor of journals should not be used for evaluating research“, *British Medical Journal* 314, S. 497-502.
- Seglen, P.O. (1994): „Causal relationship between article citedness and journal impact“, *Journal of the American Society for Information Science* 45, S. 1-11.
- Süß, S. (2007): „Grenzen einer leistungsorientierten Anreizgestaltung für Hochschullehrer“, *Hochschulmanagement* 2, S. 68-72.
- Voeth, M./Gawantka, A./Chatzopoulou, G. (2006): „Impact auf die deutsche Marketingforschung“, *Marketing Zeitschrift für Forschung und Praxis* 28, S. 7-20.
- Weingart, P. (2005): „Impact of bibliometrics upon the science system: Inadvertent consequences?“, *Scientometrics* 62, S. 117-131.
- Weingart, P./Winterhager, M. (1984): „Die Vermessung der Forschung: Theorie und Praxis der Wissenschaftsindikatoren“, Frankfurt a. M./New York: Campus Verlag.



Bisher erschienen:

## Diskussionspapiere des Instituts für Organisationsökonomik

- DP-IO 1/2012** Die Zitationshäufigkeit als Qualitätsindikator im Rahmen der Forschungsleistungsmessung  
*Harry Müller*  
Januar 2012
- DP-IO 12/2011** Ein Forschungsleistungsranking auf der Grundlage von Google Scholar  
*Alexander Dilger/Harry Müller*  
Dezember 2011
- DP-IO 11/2011** Besonderheiten der Bewerbung um Promotionsstellen und -gelegenheiten  
*Alexander Dilger*  
November 2011
- DP-IO 10/2011** 1. Jahresbericht des Instituts für Organisationsökonomik  
*Alexander Dilger/Stephanie Kiefer/Katrin Scharfenkamp*  
Oktober 2011
- DP-IO 9/2011** Corporate Governance and Employee Power in the Boardroom  
An Applied Game Theoretical Analysis  
*Benjamin Balsmeier/Andreas Bermig/Alexander Dilger/Hannah Geyer*  
September 2011
- DP-IO 8/2011** Ein Ranking von Hochschulen und (Bundes-)Ländern  
am Beispiel der Betriebswirtschaftslehre  
*Harry Müller/Alexander Dilger*  
August 2011
- DP-IO 7/2011** Befragung der Kommission Hochschulmanagement zu VHB-JOURQUAL  
*Alexander Dilger*  
Juli 2011
- DP-IO 6/2011** Director Interlocks and Executive Turnover in German Public Corporations  
A Hazard Analysis for the Period from 1996 to 2008  
*Benjamin Balsmeier/Achim Buchwald/Alexander Dilger/Jörg Lingens*  
Juni 2011
- DP-IO 5/2011** Personalökonomik  
Stärken, Schwächen und ihr Platz in der Personalwirtschaftslehre  
*Alexander Dilger*  
Mai 2011
- DP-IO 4/2011** Familienbewusste Personalpolitik und Unternehmenserfolg  
Eine empirische Untersuchung  
*Christian Lehmann*  
April 2011
- DP-IO 3/2011** Welche Unternehmen berufen Vorstandsvorsitzende und andere Vorstände als externe  
Kontrolleure?  
Eine empirische Analyse der Präsenz von externen Vorständen in den Aufsichtsräten  
deutscher Großunternehmen  
*Achim Buchwald*  
März 2011

- DP-IO 2/2011** Hat Julia aufgrund ihres Vornamens Wettbewerbsvorteile gegenüber Ayse und Chantal? Ein Experiment auf dem Beziehungs-, Nachhilfe- und Wohnungsmarkt  
*Laura Lütkenhöner*  
Februar 2011
- DP-IO 1/2011** Die dunkle Seite der Gerechtigkeit  
*Alexander Dilger*  
Januar 2011
- DP-IO 3/2010** On the Overconfidence-Effect in Teams  
*Hanke Wickhorst*  
Dezember 2010
- DP-IO 2/2010:** Leistung, Identifikation oder die Unsicherheit über den Spielausgang – was zählt wirklich? Relevante Einflussfaktoren auf die Zuschauerzahlen in der Basketball-Bundesliga  
*Hannah Geyer*  
November 2010
- DP-IO 1/2010:** A Citation Based Ranking of German-speaking Researchers in Business Administration with Data of Google Scholar  
*Alexander Dilger/Harry Müller*  
Oktober 2010



Herausgeber:  
Prof. Dr. Alexander Dilger  
Westfälische Wilhelms-Universität Münster  
Institut für Organisationsökonomik  
Scharnhorststr. 100  
D-48151 Münster

Tel: +49-251/83-24303  
Fax: +49-251/83-28429

[www.wiwi.uni-muenster.de/io](http://www.wiwi.uni-muenster.de/io)

