
Themenfeld 10: Umgang mit Zeitschriften- und Personenrankings

Bekanntlich führt der VHB selber ein Zeitschriftenranking durch, lehnt aber ein Personenranking, wie es nun zweimal durch das Handelsblatt publiziert worden ist, ab – im Unterschied beispielsweise zum Verein für Socialpolitik (dem auch VHB-Mitglieder angehören), der das Handelsblatt-Ranking sogar finanziell unterstützt. Der VHB-Vorstand hat in der Vergangenheit auch im Zusammenhang mit JOURQUAL Empfehlungen zum Umgang mit dem Zeitschriftenranking gegeben. Wesentlich zielen diese Empfehlungen auf die Art der Nutzung bei Bewerbungsverfahren für Professuren und hinsichtlich des Stellenwertes, der der Anzahl von Publikationen im Verhältnis zu anderen Bewertungskriterien (z.B. der Publikation von Lehrbüchern gegeben wird. Die Frage stellt sich aber, wie speziell die an Berufungsverfahren beteiligten Betriebswirte mit solchen Empfehlungen umgehen können – was sich auch in diesem Zusammenhang als „Gute fachliche Praxis“ erwiesen hat.

Rankings stellen eine Reihung von Objekten oder Personen im Hinblick auf einen bestimmten Zweck auf. Berufslisten für eine ausgeschriebene Professur sind solch ein Beispiel. Personenrankings gehören deshalb seit langem zum Alltag wissenschaftlicher Hochschulen

Seit etwa 25 Jahren hat die Zahl an Rankings jedoch stark zugenommen. Für den Umgang mit Rankings in der Betriebswirtschaftslehre sind hier auch die Hochschul- und Zeitschriftenrankings besonders relevant. Letztere, so das vom VHB geförderte JOURQUAL, werden unter anderem (von Dritten) dazu benutzt, die wissenschaftliche Leistung der in den gerankten Zeitschriften publizierenden Forscher und Forscherinnen zu bewerten, etwa bei Berufungsverfahren. Einige öffentliche Personenrankings, prominent dasjenige des *Handelsblattes*, fußen als Bewertungsmaßstab sogar wesentlich nur auf Zeitschriftenrankings.

In jüngerer Zeit nimmt die Kritik an den Rankings zu. Dabei ist zu unterscheiden,

- (1) ob die Konstruktion einzelner Rankings bemängelt wird
- (2) ob der Umgang mit Rankings kritisiert wird
- (3) oder ob die Nutzung von Rankings generell abgelehnt wird.

1. Konstruktion von Rankings

Unter einem *Ranking* wird im mathematischen Sinn eine transitive, reflexive und vollständige Ordnungsrelation auf einer Menge von Objekten (oder Personen) verstanden. Alle Objekte, die gleich gerankt sind, bilden eine Äquivalenzklasse. Benennt man diese Klassen mit Symbolen, welche eine natürliche oder definierte Reihenfolge besitzen (z.B. A+, A, B, C, ... oder 5*, 4*, 3*, 2*, 1*), so handelt es sich speziell um ein *Rating*¹. Typisch dabei ist oft, dass die einzelnen Ratingklassen durch exogen festgelegte Anspruchsniveaus definiert sind, welche nicht nur eine relative Ordnung innerhalb der Menge (aktuell) verglichener Objekte herstellen, sondern darüber hinaus sogar eine absolute Einstufung im Hinblick auf andere, (noch) nicht betrachtete Objekte erlauben (sollen).

¹ Ratings sind also spezielle Rankings. Deshalb wird nachfolgend vereinfachend meistens allgemein nur von Rankings gesprochen.

Soweit stellen Rankings und Ratings lediglich eine *ordinale* Bewertung dar. Sie sagen nichts über die Abstände zwischen unterschiedlich gerankten Objekten bzw. zwischen den Ratingklassen aus. Um – im Falle von Personenrankings – nur geringe Unterschiede zwischen zwei Platzierten im Vergleich mit den anderen auf einer Berufsliste zu signalisieren, erhalten diese deshalb üblicherweise dieselbe Platznummer und werden nur durch Buchstaben differenziert (z.B. Listenplätze 2a und 2b). Das vom Verein für Socialpolitik unterstützte und im Handelsblatt erscheinende Ranking (volks- bzw.) betriebswirtschaftlicher Forscher/innen und Fachbereiche geht darüber hinaus, in dem es durch die Addition von Teilwerten für die einzelnen Aufsätze in Zeitschriften Punktgesamtwerte für die wissenschaftliche Leistung der Forscher/innen bzw. Fachbereiche berechnet und somit eine *kardinale* Bewertung vornimmt. Dadurch, dass es sich sogar um eine Verhältnisskala handelt, impliziert das Verfahren Aussagen der Art, dass ein/e Forscher/in mit 20 Punkten eine doppelt so hohe Leistung erbracht habe als einer mit 10 Punkten.

Aus der Theorie mehrfacher Zielsetzungen bzw. der multiattributiven Werttheorie sind allerdings die starken Prämissen additiver Wertfunktionen bekannt, welche solchen Verfahren zu Grunde liegen. Wohl deshalb sind Punktbewertungsverfahren in Gruppenentscheidungsprozessen, z.B. bei Berufungskommissionen, kaum üblich. Sie wären sehr aufwändig, die Ziele müssten u. a. vollständig, präferenzunabhängig und messbar formuliert werden. Zudem müssten die einzelnen Gruppenmitglieder ihre detaillierten Präferenzen offenbaren und dürften nicht manipulieren.

Punktbewertungsverfahren werden bei veröffentlichten Rankings dennoch gerne eingesetzt, weil sie helfen, den Bewertungsprozess transparenter und intersubjektiv vergleichbarer zu gestalten. So erhalten etwa beim Personenranking des Handelsblatts Aufsätze in Top-Zeitschriften den Wert Eins und solche in geringer eingeschätzten Zeitschriften Werte zwischen Null und Eins. Dies impliziert, dass fünf Aufsätze in Zeitschriften mit dem Score 0,2 genau so viel wert sind wie einer in einem der Top-Journale. Es ist nicht unumstritten, ob ein solcher Trade-off sinnvoll ist. Nicht selten wird sogar die Auffassung einer lexikografischen Präferenzordnung vertreten, wonach nur Aufsätze in den Top-Zeitschriften zählen. Bei der VWL scheint es dabei einen weitgehenden Konsens über die fünf besten Journale zu geben, bei der BWL eher nicht.

Zeitschriftenrankings

Das neue VHB-JOURQUAL 3 stellt aus den genannten Gründen geringere Informationsanforderungen an die Befragung der VHB-Mitglieder, indem die jeweils bekannten Zeitschriften lediglich ordinal in die bekannten Rating-Kategorien A+, A, B, C, ... eingestuft werden. Die Gesamtbeurteilung einer Zeitschrift ergibt sich dann aus dem Median der Einzelurteile, soweit genug Urteile zusammen kommen, welche eine statistisch gesicherte Aussagekraft gewährleisten und die Manipulationsgefahr gering halten. Auch ohne bewusste Manipulationen gibt es jedoch mindestens drei kritische Punkte, die bei der Konstruktion von sowie beim Umgang mit Zeitschriftenrankings zu beachten sind oder die je nach subjektiver Einschätzung sogar eine generelle Ablehnung begründen können:

- Hat das *Journal of Finance* eine höhere Qualität als die Zeitschrift *Nature*, oder *Econometrica* als *AMJ*? Die allgemeine Beurteilung der „Qualität“ einer wissenschaftlichen Zeitschrift ist fragwürdig, weil inhaltsleer oder uneindeutig. Qualität muss inhaltlich präzise definiert werden. Für die Bedeutung einer Zeitschrift im VHB-JOURQUAL sollte es darum gehen zu klären, inwieweit die betreffende Zeitschrift zum Fortschritt der BWL als wissenschaftliche Disziplin beiträgt. Deshalb wird im neuen JOURQUAL 3 explizit danach gefragt. Wenn dann in ihren Disziplinen hoch gerankte Zeitschriften der VWL, Mathematik oder Sozial- und Naturwissenschaften nur eine geringere Bewertung erhalten sollten als *AMJ* oder *Management Science* oder gar nicht erst in das Ranking aufgenommen werden, stellt das keine generelle Geringschätzung der fachfremden Journale dar, sondern beschreibt lediglich die Relevanz der dort publizierten Erkenntnisse für die BWL, nicht für die Wissenschaft generell und keinesfalls für die betreffende Disziplin. (Dies hatten einige Mitglieder des *Vereins für Socialpolitik* missverstanden, als VWL-Journale im BWL-Ranking des VHB-JOURQUAL niedriger eingestuft wurden.) Wie anfangs vermerkt, sind Rankings nämlich immer zweckbezogen. Demnach kann die Reihung von Objekten je nach dem verfolgten Zweck verschieden ausfallen. Von daher ist es problematisch, wenn im Handelsblattranking der BWL VWL-Zeitschriften mit ihrer VWL-Einstufung eingehen, ohne ihre BWL-Relevanz zu bestimmen.
- Zeitschriften, welche Erkenntnisse aus zwei oder mehr Disziplinen veröffentlichen, können daher je nach Disziplin unterschiedlich bewertet werden. Weil die BWL selber wieder zum Teil sehr heterogene Teilgebiete umfasst, deren Erkenntnisziele und wissenschaftliche Methodik breit variieren, kann das Ranking

von Zeitschriften in den Überlappungsbereichen solchermaßen differierender Teilgebiete voneinander abweichen, z.B. beim Financial Accounting zwischen den Forschern der Finanzwirtschaft einerseits und denen des Rechnungswesens andererseits. Es ist prinzipiell unmöglich, ein nicht-triviales (eindimensionales) Ranking von Zeitschriften aufzustellen, das einerseits deren Bedeutung für die BWL insgesamt erfasst und andererseits gleichzeitig eine individuelle Bewertung je nach Teilgebiet zulässt. Dieses Dilemma kann sogar auch innerhalb eines Teilgebiets auftreten, wenn deren Forscher verschiedenen Paradigmata folgen. So gesehen gibt ein Ranking über die Bedeutung einer Zeitschrift (für eine Disziplin insgesamt oder aber für eines ihrer Teilgebiete) lediglich den über alle befragten Forscher/innen gemittelten Gesamteindruck wieder. Streuungsmaße können helfen, die Validität und Aussagekraft eines Ranking besser einschätzen zu können.

- Alternativ zur unmittelbaren subjektiven Beurteilung durch Forscher/innen werden die Impact-Faktoren von Zeitschriften als Rankingmaße verwendet. Sie haben allerdings ebenfalls Nachteile, zum Teil ähnliche. So sind Manipulationen durch Zitationskartelle nicht ausgeschlossen, wenn auch wohl weniger relevant. Auch variiert die Zitationskultur verschiedener Forschercommunities teilweise so stark, dass interdisziplinäre Leistungsvergleiche nicht angemessen sind. Der (im *Web of Knowledge*) üblicherweise auf nur zwei Jahre beschränkte Betrachtungszeitraum ausgewerteter Zitate kann darüber hinaus die Bedeutung solcher viel zitierter Publikationen nicht adäquat erfassen, deren Relevanz erst später erkannt wird und die dann eventuell völlig neue Forschungsgebiete eröffnen. Darüber hinaus ist die Verteilung der Zitationen auf die in einer Zeitschrift erschienenen Artikel extrem schief und abhängig vom Artikeltyp. So werden Übersichtsaufsätze häufiger zitiert, obwohl sie nicht unbedingt innovativ sind. Und viele Artikel werden gar nicht oder sehr selten zitiert, während wenige Artikel fast alle Zitate auf sich vereinigen.

Personenrankings

Wegen der schiefen Verteilung der Zitate auf die Zeitschriftenartikel ist es logisch inkonsistent, wenn ein Forscherranking sich auf die Impact-Faktoren von Zeitschriften beruft und die Aufsätze damit gewichtet. Wenn man schon Zitate als Ausdruck der wissenschaftlichen Bedeutung eines Aufsatzes begreift (was mit guten Gründen in Frage gestellt werden kann), dann wären unmittelbare, autorenbezogene Zitationsanalysen viel sinnvoller. Sie sind mittels des *Web of Knowledge* oder mittels *Google Scholar* durchaus möglich, wenn auch nicht problemlos. Mit *Google Scholar* werden zudem Bücher und deren Zitationen erfasst, bislang allerdings hauptsächlich nur englischsprachige.

Personenrankings mittels Zitationsanalysen scheinen kaum üblich zu sein, vor allem wohl auch deshalb, weil die Zitationen deutscher Betriebswirte in internationalen Literaturdatenbanken ebenfalls extrem schief verteilt sind. Wenn für Forscher/innen nur wenige Zitate existieren – wie es für viele, insbesondere junge Forscher/innen der Fall ist –, sind Zitationsanalysen wenig aussagekräftig und kaum valide, um die jeweilige Forschungsleistung aufzuzeigen. Für ältere Forscher/innen bildet dagegen der *Hirsch-Index h* einen robusten und zuverlässigen Indikator, der eine untere Schranke für die kombinierte Publikations- und Zitationshäufigkeit bestimmt. So besagt $h = 20$, dass die Person (oder Fakultät) zwanzig Aufsätze in der zu Grunde liegenden Literaturdatenbank aufweist, die mindestens zwanzig Mal zitiert sind, und dass somit mindestens 400 Zitationen auf die Person (bzw. die Fakultät) entfallen.

Auf einzelnen Forschungsleistungen basierende Personenrankings verführen insbesondere zu Werturteilen, höher gerankte Professoren/innen seien besser als niedriger gerankte Kollegen/innen. Allerdings muss man sich bewusst sein, dass vor allem auf Literaturdatenbanken basierende Auswertungen letztendlich immer nur einen Ausschnitt professoraler (Gesamt-)Leistungen darstellen, nämlich genau solche, die von den Datenbanken entsprechend erfasst werden. Mangelnde Erfassungen zeigen sich letztendlich auch in oben angesprochenen schiefen Verteilungen, vor allem in Bezug auf Zitationen. Zusätzlich ist das Bild eines Professors bzw. einer Professorin durch unterschiedliche Aufgabenbestandteile geprägt, die in solchen Rankings letztendlich nie berücksichtigt werden. So fokussiert jedes Ranking zwangsläufig stets auf einen Ausschnitt aus dem totalen professoralen Leistungsspektrum, was aber zumeist kaum thematisiert wird.

Bei Berufungsverfahren wird kein vollständiges, striktes Ranking aller Bewerberinnen und Bewerber erstellt. Für den Zweck der letztendlichen Berufung einer einzigen Kandidatin oder eines einzigen Kandidaten reicht in der Regel ein striktes Ranking von nur drei Personen auf einer Berufsungsliste aus. Diese werden normalerweise dadurch ermittelt, dass zu Beginn Ratingklassen gebildet werden, in die alle Bewerberinnen und Bewerber in einem ersten Screening-Prozess eingeteilt werden. Vielfach üblich sind folgende drei Klassen:

(A) mit Priorität näher zu sichten, weil offensichtlich die Anforderungen erfüllend, (B) gegebenenfalls auch noch zu prüfen, aber nach „Papierform“ nur bedingt geeignet, (C) anscheinend ungeeignet oder zumindest deutlich schlechter als die Personen in den beiden anderen Klassen (sofern diese schon ausreichend viele Personen enthalten). Im weiteren Verlauf des Verfahrens kommen dann weitere Bewertungsaspekte ins Spiel, so der persönliche Eindruck im Probevortrag.

2. Umgang mit Rankings

Persönliche langjährige Erfahrungen zeigen, dass bei Berufungsverfahren „harte“ Kriterien, wie nachweisliche Publikationen in allseits anerkannten, hochrangigen Zeitschriften, heute eine wichtige, allerdings ambivalente Rolle spielen: Einerseits helfen sie der Kommission, die Forschungsleistung besser einzuschätzen – vor allem in kleinen Fachbereichen mit wenig Expertise im ausgeschriebenen Fachgebiet – und verhindern so Fehleinschätzungen oder Manipulationen zu Gunsten weniger qualifizierter Personen. Andererseits verengen sie das Profil der schließlich ausgewählten Forscher auf die in den anerkannten Journals akzeptierten Fachgebiete und Paradigmen. Weil diese anerkannten Journale in der Regel in den Kernbereichen der Disziplin liegen und normalerweise dem herrschenden Paradigma folgen, besteht die große Gefahr, dass dadurch innovative Erkenntnisse und neuartige, noch unerprobte Konzepte sowie Forschung in Grenzgebieten vernachlässigt werden.

Die Einführung des Zeitschriftenrankings VHB-JOURQUAL in 2003 hat die Bewertung der Forschungsleistung deutscher Betriebswirte stark erleichtert und kann dabei sowohl zu besseren als auch zu schlechteren Resultaten führen. Einerseits wird dadurch die Gefahr verringert, schwache Forscher/innen auf Professuren zu berufen. Andererseits wird die Tendenz zu einer gewissen Fehllesung gefördert, wenn Berufungen allein mittels Abzählung der Publikationen in hoch gerankten Zeitschriften erfolgen und dabei andere, oft wichtigere Aspekte ignoriert werden (z.B. die fachliche Ausrichtung oder die Lehrqualifikation). Fakultäten bilden Teams mit (nichtlinearen) Synergieeffekten, sodass eine Berufung eher der Suche nach einem passenden Mosaikstein ähnelt, weshalb die einzelnen Qualifikationsanforderungen situationsgerecht zu bewerten sind, also nicht unbedingt nach einem für alle identischen Standard: Wenn der Torwart für eine Fußballmannschaft gesucht wird, spielen Stürmereigenschaften eine geringe Rolle.

Neben dem Einsatz von Zeitschriftenrankings zur (ersten) Bewerberauswahl im Rahmen von Berufungsverfahren sind auch weitere Verwendungsmöglichkeiten kritisch zu reflektieren. So können Zeitschriftenrankings auch die Auswahl eines geeigneten Organs zur Publikation wissenschaftlicher Erkenntnisse unterstützen. Insbesondere wird dem wissenschaftlichen Nachwuchs aufgezeigt, wie stark oder schwach einzelne Zeitschriften bzw. der dort publizierte Content derzeit von der für sie relevanten Scientific Community eingeschätzt werden. Nachteilig könnte sich bei einer derartigen Verwendung jedoch auswirken, dass es bei der Auswahl eines Publikationsorgans nicht mehr darum geht, ob die Zeitschrift auch wirklich inhaltlich zu den erforschten Ergebnissen passt. Vielmehr könnte man nunmehr gewillt sein, die Publikationsmedien nach der Ranking-Bewertung bzw. der Ratingklasse auszuwählen. Die Problematik verschärft sich vor allem, da es Teilgebiete in der BWL gibt, bei denen viele Zeitschriften eine hohe Bewertung erhalten und andere, die keine Zeitschriften in den höchsten Bewertungskategorien aufweisen.

Rankings erzwingen eine Aggregation der verschiedenen Aspekte, welche für die Bewertung der zu rankenden Objekte oder Personen relevant sind. Vor allem solche Rankings, welche mittels Punktbewertungsverfahren ermittelt werden, stellen hohe Anforderungen. Es sei denn, es würde ein allgemein akzeptierter Maßstab existieren (quasi als „Währungseinheit“), welcher eine transparente, intersubjektiv nachprüfbare Bestimmung der Trade-offs bei der Abwägung der Vor- und Nachteile in den verschiedenen relevanten Gesichtspunkten ermöglicht. Bei der Bewertung von Forschung und Lehre ist das jedoch nicht der Fall, u.a. mangels adäquater Märkte für diese Leistungen. Deshalb bildet jedes Ranking letztlich nur das Ordnungskriterium ab, welches zu der konkreten Reihung der Objekte bzw. Personen geführt hat, nicht mehr, aber auch nicht weniger. So besagt eine Berufsungsliste nicht zwingend, dass der Erstplatzierte am besten für die ausgeschriebene Stelle geeignet ist, sondern unter günstigen Umständen nur, dass die Mehrheit der Berufungskommissionsmitglieder ihn dafür hält, gegebenenfalls aber nicht einmal das. Und ein Personenranking wie das vom Handelsblatt gibt zu-nächst einmal nur an, wie viele Punkte nach dem zu Grunde gelegten Berechnungsschema die betreffende Person in den als relevant erachteten Zeitschriften erreicht hat. Es mag eine positive, nicht vernachlässigbare Korrelation zwischen dieser Punktzahl und der „wahren“ Forschungsleistung der gerankten Personen geben. Es ist jedoch gewagt, zumindest unwissenschaftlich, aus

der Handelsblattpunktzahl eines einzelnen Forschers oder einer einzelnen Forscherin definitiv auf seine oder ihre gesamte Forschungsleistung zu schließen.

3. Ablehnung von Rankings

Wie jede Kennzahl im Controlling bieten Punktzahlen zur Aufstellung von Rankings Vor- und Nachteile. Sie stehen in dem Spannungsverhältnis der beiden bekannten Praktikerregeln: „You can't manage what you can't measure!“ und „What you measure is what you get!“ Anders gesagt geben sie ebenso wie die damit gebildeten Rankings und die ihnen zugrunde liegenden Verfahren den jeweiligen Nutzern (z.B. Berufungskommission, Hochschulleitung, Staat, Politik, Öffentlichkeit) Instrumente in die Hand, um damit bestimmte Ziele zu erreichen. Wie alle Instrumente können sie (absichtlich oder unabsichtlich) falsch genutzt oder sogar missbraucht werden. Wenn die negativen Folgen die positiven überwiegen, sollte man die Abschaffung solcher Instrumente vorsehen. Der VHB fördert deshalb keine Personenrankings. Beim Ranking bzw. Rating wissenschaftlicher BWL-Zeitschriften sieht die überwiegende Mehrheit der VHB-Mitglieder anscheinend (noch) mehr Vor- als Nachteile, sofern den Empfehlungen der damaligen VHB-Vorsitzenden zur verantwortlichen und reflektierten Nutzung des JOURQUAL 2.1 gefolgt wird. Das JOURQUAL 3 greift wesentliche Kritikpunkte auf und möchte dadurch die Nutzungsmöglichkeiten verbessern und Anwendungsgefahren mindern. Die vorangehenden Überlegungen sollen gute fachliche Praktiken im Umgang mit Rankings unterstützen. Da es bei der Frage der zukünftigen Auswirkungen solcher Rankings auf die wissenschaftliche Landschaft jedoch um subjektive Prognosen und Risikoeinstellungen geht, ist andererseits auch die generelle Ablehnung von Rankings, insbesondere wegen der Gefahr irreführender Anreizwirkungen, rational.

Verband der Hochschullehrer für Betriebswirtschaft e.V.
Verbandsgeschäftsführerin: Tina Osteneck
Geschäftsstelle: Reitstallstr. 7 – 37073 Göttingen – Deutschland
Tel.: +49 (0)551 – 797 78 566, Fax: +49 (0)551 – 797 78 567
E-Mail: info@vhbonline.org – URL: <http://vhbonline.org>