



WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER



Diskussionspapier des Instituts für Organisationsökonomik

8/2017

Zur Stärkung wissenschaftlicher Kritik

Alexander Dilger

Discussion Paper of the
Institute for Organisational Economics

**Diskussionspapier des
Instituts für Organisationsökonomik
8/2017**

August 2017

ISSN 2191-2475

Zur Stärkung wissenschaftlicher Kritik

Alexander Dilger

Zusammenfassung

Kritik ist ein wichtiger Teil der Wissenschaft, der jedoch nicht immer hinreichend gewürdigt wird. Dieser Beitrag zeigt die Bedeutung wissenschaftlicher einschließlich destruktiver Kritik und das Fehlen von Anreizen dazu auf und beinhaltet konstruktive Verbesserungsvorschläge.

JEL-Codes: B40, D70, D80, I23, M51

On Strengthening of Academic Critique

Abstract

Critique is an important part of science and academia that is not always sufficiently recognised. This paper shows the importance of academic including destructive criticism and the missing of incentives for it and contains constructive suggestions for improvement.

Im Internet unter:

http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO_08_2017

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

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

Zur Stärkung wissenschaftlicher Kritik

1. Einleitung

Wissenschaft ist per definitionem kritisch.¹ Ohne Kritik verkümmert Wissenschaft zu Scholastik oder sogar reiner Dogmatik. Allerdings ist ihr kritischer Charakter stets gefährdet, da es nicht genügend individuelle (oder auch kollektive²) Anreize für begründete, insbesondere auch destruktive Kritik gibt. So droht mangels Kritik ein Epizyklus an den nächsten gefügt zu werden, anstatt dass grundlegend neue Gedanken entwickelt und die besten davon herausgefiltert werden.

Im folgenden, zweiten Abschnitt wird die Bedeutung von sachlicher Kritik in der Wissenschaft herausgestellt, selbst wenn sie destruktiv sein sollte und vermeintliche Erkenntnisse begründet zerstört. Im dritten Abschnitt wird gezeigt, dass es nicht genügend Anreize zur Generierung solcher Kritik gibt und ehrliche Kritiker sogar mit großer Wahrscheinlichkeit ausgesiekt werden. Der vierte Abschnitt behandelt einige Reformmöglichkeiten, um gerechtfertigte Kritik zu erleichtern und gute Kritiker zu belohnen. Schließlich beinhaltet der fünfte Abschnitt ein kurzes Fazit.

2. Bedeutung wissenschaftlicher Kritik

Konstruktive Kritik mag in der Wissenschaft (und auch darüber hinaus) noch willkommen sein, destruktive Kritik ist es in den seltensten Fällen. Dies gilt nicht nur für den Kritisierten, sondern für die ganze wissenschaftliche Gemeinschaft. Da destruktive Kritik wissenschaftliche Erkenntnisse angreift und zu zerstören versucht, scheint sie dem wissenschaftlichen Fortschritt entgegenzulaufen. Nach erfolgreicher Destruktion ist der akzeptierte Wissensbestand kleiner als zuvor, bei erfolglosem Destruktionsversuch nicht größer. Allerdings ist dieser Eindruck selbst eine Täuschung, die es kritisch zu destruieren gilt.

Die destruktive Kritik befördert die Wissenschaft und ihren Fortschritt, indem sie vermeintliches Wissen als solches, nämlich vermeintlich, offenbart. Irrtümer werden als Irrtümer erkannt und nicht mehr für richtig gehalten, was einen Erkenntnisfortschritt darstellt. Der Be-

¹ Das gilt auch für die moderne Philosophie, siehe z. B. die „Kritik der reinen Vernunft“ von Kant (1998).

² So argumentieren Armbrüster/Kieser (2001), dass u. a. deshalb der Rat von Unternehmensberater eher nachgefragt wird als der von Wissenschaftlern, weil erstere sich im Gegensatz zu letzteren nicht gegenseitig kritisieren und widersprechen würden. Aus Marketinggesichtspunkten ist Kritik also weniger empfehlenswert, trotzdem jedoch für die Wissenschaft unverzichtbar.

stand der anerkannten Wahrheiten schrumpft,³ wird aber zugleich in seiner Substanz besser. Hinzu kommt das Wissen um die Irrtümer, die zumindest in derselben Weise dadurch in Zukunft nicht mehr auftreten sollten. Es wird Platz geschaffen für neue, konstruktive Erkenntnisse, die sonst vermutlich aus falschen Gründen abgelehnt oder erst gar nicht gesucht worden wären.

Destruktive Kritik hat daneben den Vorteil, dass ihre Richtigkeit in aller Regel leichter beurteilt werden kann als die von konstruktiven Erkenntnissen. Dies führt schnell zu dem Vorurteil, dass sie auch leichter zu produzieren sei als Konstruktives, was auf sorgfältige Destruktion jedoch keinesfalls zutreffen muss. Natürlich lässt sich leicht Destruktives äußern und behaupten, doch das gilt für Konstruktives gleichermaßen. Entscheidend ist das Begründen und Belegen der Aussagen, was sowohl auf konstruktive als auch auf destruktive Aussagen zutrifft. Es kann im destruktiven Fall ebenso schwer sein wie im konstruktiven oder noch schwerer, doch wenn es einmal gelungen ist, ist es häufig logisch zwingend.⁴

Das schließt nicht aus, dass kritische Aussagen ihrerseits wieder der Kritik unterzogen werden können. Außerdem ist zu bedenken, dass die Fähigkeiten der Menschen ungleich verteilt und ausgebildet sind, so dass für den einen Konstruktives, für den anderen Destruktives leichter zu leisten ist. Das sollte im Rahmen wissenschaftlicher Arbeitsteilung genutzt werden, statt nur konstruktive Artikel für den Beginn einer wissenschaftlichen Karriere zuzulassen, um danach jedem Wissenschaftler das Verfassen kritischer Gutachten zuzumuten.

Destruktive Kritik hat, wenn sie berechtigt ist, den weiteren Vorteil, dass sie schlechte Wissenschaft und in manchen Fällen auch schlechte Wissenschaftler als solche aufdeckt. Dadurch können die knappen Ressourcen, sowohl materieller Art, etwa Stellen und Ausstattung, als auch immaterielle Ressourcen, z. B. Aufmerksamkeit und Prestige, besseren Zwecken zugeführt werden. Dabei wirkt die Kritik nicht nur ex post selektierend, sondern ihre Möglichkeit steigert bereits ex ante die Motivation, möglichst gute oder zumindest kritikresistente Wissenschaft zu betreiben. Der Wunsch nach Absicherung gegen Kritik kann natürlich auch übertrieben werden und Kreativität oder im Extremfall wissenschaftliche Betätigung überhaupt hemmen. Das spricht jedoch nicht gegen die Kritik, sondern für einen angemesseneren und

³ Kuhn (1962) und auch Lakatos (1977) weisen allerdings darauf hin, dass ein ganzes Paradigma nicht einfach falsifiziert und ersatzlos verworfen, sondern nur durch ein anderes Paradigma abgelöst werden kann. Das unterstreicht die ohnehin nicht strittige Bedeutung konstruktiver wissenschaftlicher Arbeit.

⁴ Die Argumentation hier ähnelt der von Popper (1994), wonach empirische Aussagen nur falsifiziert, nicht induktiv bestätigt werden können. Allerdings ist Kritik nicht nur an empirischen Aussagen möglich, wie auch deduktive Logik nicht der einzige wissenschaftliche Maßstab ist.

entspannteren Umgang mit ihr. Gerade wenn Kritik häufig ist, wird der Umgang mit ihr vertrauter.

Schwere wissenschaftliche Verfehlungen wie Plagiate oder Datenfälschungen sollen möglichst vollkommen unterbunden und entsprechend angeprangert und bestraft werden. Dazu ist es hilfreich, wenn Entdecker solcher Verfehlungen belohnt statt totgeschwiegen oder sogar bestraft werden. Irrtümer, insbesondere solche, die ohne Verschulden zustande kamen oder sogar unvermeidlich waren, sind hingegen einfach auszuräumen und nicht zu sanktionieren. Wenn Kritik als alltäglich und normal erlebt wird, ist die Feststellung von Fehlern keine um jeden Preis zu vermeidende, schlimmstenfalls zu verschleierte Katastrophe mehr für ihre Urheber, sondern ein begrüßenswerter Erkenntnisfortschritt.

3. Mangelnde Anreize zur Kritik

Die gegenwärtige Situation in den meisten Wissenschaften lädt nicht gerade ein zur Kritik bzw. zu deren ungeschminkter Äußerung. Institutionell gibt es einige geregelte Kritikverfahren, die jedoch alle mit gravierenden Mängeln behaftet sind. Insbesondere fehlt regelmäßig der Anreiz, ehrlich und begründet Kritik zu üben, insbesondere wenn diese negativ ausfallen müsste.

So macht sich, wer offen Kritik übt, in der Regel mindestens den Kritisierten zum Feind, häufig auch dessen Freunde, Protégés etc. Dies ist unabhängig von der Berechtigung der Kritik bzw. gerade bei berechtigter Kritik eher noch stärker der Fall. Der Grund dürfte darin zu suchen sein, dass berechtigte Kritik meist als Schwäche des Kritisierten ausgelegt wird. Um nicht als kritikwürdig zu erscheinen, wird die Berechtigung der Kritik bestritten. So entwickelt sich schnell eine inhaltliche Frage zu einem Machtkampf. Da dabei häufig beide Seiten verlieren und fast immer die schwächere, ist es eine Frage der Klugheit, auf Kritik zu verzichten, selbst wenn diese inhaltlich berechtigt sein sollte. Weil Wissenschaftlicher nicht dumm sind und in einem sehr kompetitiven Umfeld taktische Fehler die Karriere kosten können, ist mit vergleichsweise wenig offen geäußelter Kritik zu rechnen. Dies ließe sich nur durch entsprechend starke Anreize zur Kritik ändern, etwa eine deutliche Mehrung des Ruhms erfolgreicher Kritiker. Doch selbst dann bliebe die Erfolgsbeurteilung ein kritischer Punkt und könnte die Zunahme von Machtkämpfen statt sachlicher Kritik drohen.

Die große Bedeutung von Macht und Position im Verhältnis zu inhaltlichen Argumenten lässt sich gut im wissenschaftlichen Qualifizierungsprozess beobachten. Dort spielt Kritik, häufig

auch destruktive Kritik, durchaus eine Rolle, jedoch nicht immer eine glückliche. Die Kritik wird in der Regel offen geäußert, die Identität des Kritikers ist bekannt, doch hat dieser eine höhere Position und in der Regel erhebliche Macht über den Kritisierten. Es fängt an bei Klausuren und mündlichen Prüfungen von Studierenden. Die Prüfer haben den Prüfungsstoff mehr oder weniger gut vermittelt und überprüfen danach selbst den Vermittlungserfolg. Die Studierenden haben kaum Mittel, sich gegen ungerechtfertigte Kritik an dieser Stelle zu wehren, während gegen ungerechtfertigtes Lob ohnehin keiner wird vorgehen wollen. Dies kann dazu führen, dass völlig sachfremde Kriterien über den Prüfungserfolg entscheiden, z. B. die Kleidung, persönliche Sympathie oder Mitleid. Kritik in die Gegenrichtung, also am Prüfer durch den Prüfling, ist fast nie zu empfehlen, selbst bzw. gerade wenn ihre inhaltliche Berechtigung offensichtlich sein sollte. Die Qualität der Prüfungen und der darin geäußerten Kritik hängt im Wesentlichen von der Person des Prüfers ab, der seinerseits nicht nach seinen Kritikerfähigkeiten als Prüfer ausgewählt wurde, sondern z. B. für das Verfassen konstruktiver wissenschaftlicher Beiträge und natürlich davor sein erfolgreiches Überstehen vieler Prüfungen in der Rolle des Geprüften.

Auf späteren Stufen des Qualifizierungsprozesses, also bei Promotionen und insbesondere Habilitationen sowie Juniorprofessuren, nimmt die Bedeutung der Kollegialprüfung zu, wenngleich diese zunehmend durch kumulative Dissertationen und Habilitationen an Gutachter von Zeitschriften outgesourct wird. Wenn bei normalen Studenten Zweitprüfer und -gutachten vorgeschrieben sind, so wird das Ergebnis des Erstprüfers meistens ohne inhaltliche Kontrolle durch den Zweitkorrektor übernommen, was die Position des Studierenden eher noch schwächt, da nur einer ihn beurteilt hat, zwei jedoch mit ihrem Namen für das Urteil einstehen. Bei Promotionen und Habilitationen werden hingegen die weiteren Gutachten in der Regel ernster genommen und können durchaus abweichen. Der Kandidat muss dann nicht nur einen Gutachter überzeugen, sondern mehrere. Potentiell könnte das die Sachlichkeit des Verfahrens erhöhen, faktisch bringt es weitere Unwägbarkeiten mit sich und macht das Verfahren sehr „politisch“. So versucht regelmäßig der Erstgutachter, der die Arbeit angeleitet und den Kandidaten über Jahren hinweg betreut oder sogar beschäftigt hat, das Verfahren zu einem erfolgreichen Abschluss zu bringen, während die weiteren Gutachter je nach Stellung zu ihm, also dem Erstgutachter, dieses Vorhaben zu unterstützen oder torpedieren versuchen. Bei Habilitationen und Juniorprofessuren stimmt schließlich die ganze Fakultät ab. Kritik ist gefährlich, wenn der Betreuer stark ist. Umgekehrt kann bei einem sehr schwachen und unbeliebten Betreuer das Unterlassen von Kritik durch die Kollegen bereits negativ ausgelegt werden. Um seine Stärke zu demonstrieren, wird der Betreuer als der mit der Arbeit am besten

Vertraute im Übrigen meistens seine eigene inhaltliche Kritik zurückstellen und auch schwache Kandidaten durchzubringen versuchen. Juniorprofessoren dürfte es schließlich zum Nachteil gereichen, dass sie im formellen Sinne gar keinen Vorgesetzten und Betreuer haben, der sich bei den Kollegen für sie einsetzt.

Ähnlich politisch verlaufen die institutionalisierten Diskussionen auf Tagungen und in Sammelbänden. Kritik ist hier möglich bis formal erwünscht, jedoch nur in der hierarchischen Richtung von oben nach unten üblich und ratsam. Gleichrangige Kollegen werden untereinander mit Kritik, vor allem öffentlich geäußelter, sehr vorsichtig sein. Dasselbe gilt für Kritik an den Untergebenen sehr starker Kollegen. Stets dürfte weniger die Sache als mögliche Gegenmaßnahmen im Vordergrund stehen. Die Furcht vor Gegenmaßnahmen ist vernünftig, da sie die eigene Karriere beeinträchtigen können, während sich durch noch so berechtigte und erfolgreiche Kritik kaum etwas gewinnen lässt. Um Schutzbedürftige um sich zu sammeln und die eigene Macht zu stärken, ist auch das implizite Androhen und gegebenenfalls Durchführen solcher Gegenmaßnahmen rational, so dass es sich um ein Gleichgewicht handelt, Kritik nach Macht- statt Sachgesichtspunkten (nicht) zu äußern.

Es stellt sich die Frage, warum hier dann doch gelegentlich durchaus Kritik zu vernehmen ist, z. T. sehr verletzend bis vernichtend. Im Einzelfall können es tatsächlich sachliche Gründe sein. So mag jemandem ein Vortrag so schlecht erscheinen, dass er meint, das unbedingt feststellen zu müssen. Das hat allerdings etwas Naives an sich, weil die behandelten Machtgründe dadurch keineswegs verschwinden, sondern eher noch zunehmen. Je mehr eine Kritik (zu)trifft, desto nötiger ist ihre Abwehr. Außerdem ist nicht ersichtlich, warum als offensichtlich Angenommenes (hier die Kritikwürdigkeit) noch geäußert werden muss. Eine Überlegung mag sein, dass die inhaltliche Kritik so stark ist, dass sie die Machtbalance verändert, den Kritisierten zu Fall und dem Kritiker großen Ruhm bringt. Dies dürfte jedoch in den wenigsten Fällen gelingen, insbesondere wenn der Kritisierte noch am Leben oder gar anwesend ist.

Anders verhält es sich, wenn der Kritiker über ein deutlich höheres Standing verfügt als der Kritisierte. Hier kann er sich die Kritik leisten, während sie den Kritisierten seine Karriere kosten kann. Trotzdem fragt sich, warum er sich die Kritik leistet. Eine Möglichkeit ist natürlich, dass er die Kritik für sachlich angemessen hält. Es reicht dann eine nüchterne Bemerkung aus, deren Tragweite gar nicht allen Anwesenden, vielleicht nicht einmal dem Kritisierten selbst bewusst zu werden braucht. Hinzutreten kann die Überlegung, den Urheber des derart Kritikwürdigen nicht ungeschoren davonkommen zu lassen. Dann ist die Kritik deut-

licher zu akzentuieren, damit sie jeder versteht. Aber dieses öffentliche Bloßstellen ist auch ohne sachlich gerechtfertigten Grund möglich, etwa um die eigene Macht zu demonstrieren, dem Betreuer des Kritisierten zu schaden oder eigenen Betreuten einen relativen Karrierevorteil zu verschaffen.

Ein weiterer häufiger Fall ist die Kritik zwischen Anhängern verschiedener wissenschaftlicher Schulen bzw. Paradigmen. Jemand wird (vermeintlich) gefahrlos kritisiert, weil er eine ganz andere Wissenschaftsauffassung vertritt, die der eigenen zuwiderläuft. Da die Kollegen, die der eigenen Schule angehören, als allein relevante Wissenschaftsgemeinschaft erscheinen und der Kritik zustimmen, ist diese (inter)subjektiv gut begründet und dem eigenen Ruf zuträglich. Der Kritisierte und die Anhänger seiner Schule werden hingegen als irrend und irrelevant erlebt. Unabhängig davon, welches Paradigma tatsächlich richtig ist, trägt solche Kritik in der Regel wenig zum wissenschaftlichen Fortschritt bei, da sie die so Kritisierten keineswegs überzeugt werden, während der Kritiker und seine Verbündeten schon vorher überzeugt waren. Am ehesten ist noch vorstellbar, dass vorher neutrale Beobachter, also vor allem Novizen in dem Fach, so etwas lernen können. Einen echten Dialog zwischen Paradigmen ersetzt das jedoch nicht. Weiterführend wäre hier eine innere Kritik, die Schwachstellen unter den Prämissen des Kritisierten aufzeigt, nicht unter gänzlich anderen Voraussetzungen. Außerdem hat natürlich auch diese Kritik ihre Karriererisiken, da auch die Angehörigen anderer Schulen vermutlich einmal über den Kritiker werden abstimmen dürfen.

Folglich bietet es sich an, nach Verfahren zu suchen, wie Kritik weniger offen geäußert und zugleich qualitativ verbessert werden kann. Als das wohl wichtigste Verfahren in dieser Hinsicht hat sich die (doppelt)blinde Begutachtung von vor allem Zeitschriftenartikeln etabliert.⁵ Kaum eine renommierte wissenschaftliche Zeitschrift kommt inzwischen ohne diese Art der Begutachtung aus. Das doppelblinde Begutachtungsverfahren wurde teilweise sogar selbst als der entscheidende Qualitätsfaktor gesehen wird, so als wären derart begutachtete Zeitschriften und die in ihnen enthaltenen Artikel per se gut, nicht doppelblind begutachtete Zeitschriften und Artikel hingegen grundsätzlich schlecht(er). In letzter Zeit ist jedoch zu beobachten, dass gerade die führenden Zeitschriften die doppelblinde Begutachtung wieder aufgeben und auf eine einfachblinde Begutachtung umstellen, bei welcher die Autoren von Artikeln die Gutachter nicht kennen, die Gutachter jedoch die Namen der Autoren erfahren, die sie bei genauer Kenntnis der relevanten Forscher und Forschung ohnehin vermutet oder sogar gewusst hätten (s. u.).

⁵ Zur historischen Entwicklung von Peer Review siehe Spier (2002).

Für das Verfahren der zumindest einfachblinden Begutachtung spricht, insbesondere bei Berücksichtigung der Alternativen, dass die Gutachter durch ihre Anonymität geschützt werden und so eher Kritik üben und nach rein inhaltlichen Kriterien entscheiden können, als wenn sie Sanktionen befürchten müssten (oder auch Belohnung erhoffen dürften, was im Prinzip auf das Gleiche hinausläuft, da dann das Ausbleiben der Belohnung gewissermaßen die Sanktion darstellen würde). Ein gravierender Nachteil ist jedoch, dass die Anonymität nicht nur berechtigte Kritik ermöglicht, sondern auch jede andere Form der Begutachtung, sei sie noch so schlecht und unsachlich.⁶ Für die Gutachter besteht kein Anreiz, sich besondere Mühe bei der Begutachtung zu geben. Sie können nach Belieben Manuskripte befürworten oder zurückweisen, mit starken, schwachen oder auch gar keinen Gründen. Im schlimmsten Fall bekommt der Herausgeber einen schlechten Eindruck von der Qualität der Gutachten, so dass er den betreffenden Gutachter weniger oft oder gar nicht mehr heranziehen wird, was jedoch durchaus in dessen Eigeninteresse sein kann, da es Arbeit einspart, die zumindest direkt nichts einbringt.

Anonyme Begutachtung erleichtert also Kritik, gibt aber keinen Anreiz, sie tatsächlich vernünftig auszuüben. Wer dies trotzdem tut, unternimmt dies auf eigene Kosten. Während ein anderer die Zeit für eigene Veröffentlichungen und damit das Voranbringen seiner Karriere aufbringt, gerät der gewissenhafte Gutachter ins Hintertreffen. Ökonomisch rational wäre es hingegen, Gutachten sehr schnell zu schreiben, nach nur oberflächlicher oder sogar unvollständiger Lektüre der Manuskripte, und die Empfehlung einer (Nicht-)Veröffentlichung nach anderen als rein sachlichen Kriterien zu geben, z. B. nach der Identität des Autors. Befreundete Forscher oder auch Anhänger der eigenen wissenschaftlichen Schule werden durchgewunken, missliebige Kollegen oder auch direkte Konkurrenten (gegebenenfalls der eigenen Schüler) auf dem Berufungsmarkt werden hingegen in jedem Falle abgelehnt. Irgendein inhaltlicher oder formaler Mangel lässt sich eigentlich immer finden, zumal bei Bewertungsspielräumen jede Eigenschaft ins Negative gewendet werden kann (entweder ist der Artikel zu innovativ oder nicht innovativ genug, vernachlässigt wichtige Quellen oder listet zu viele auf etc.). Wenn gar nichts anderes hilft, lässt sich immer noch die ganze Thematik als uninteressant oder zumindest für die jeweilige Zeitschrift nicht einschlägig genug abqualifizieren.

Die Erkenntnis von „Freund“ und „Feind“ ist auch nicht sonderlich schwer, denn selbst die offiziell doppelblinde Begutachtung, bei der weder der Autor den Gutachter noch dieser den

⁶ Zu praktischen Problemen des Begutachtungsprozesses und Lösungsmöglichkeiten siehe Mulligan (2004).

Autoren kennt, ist in aller Regel nur einfachblind,⁷ wobei selbst das nicht immer so sein muss. Die Identität des Autors ergibt sich häufig schon durch die Existenz eines namentlich gekennzeichneten Diskussionspapiers mit dem gleichen Inhalt wie das anonyme Manuskript. Experten auf dem wissenschaftlichen Teilgebiet, welches gute Gutachter eigentlich sein sollten, kennen die betreffenden Diskussionspapiere bereits, andernfalls hilft eine kurze Google-Recherche. Weiterhin nennt sich der Autor häufig selbst in seinem Manuskript („wie ich schon vor drei Jahren zeigte“ mit namentlicher Quellenangabe) oder kann aus seinem Zitierverhalten, seiner Schreibweise, den verwendeten Daten etc. erschlossen werden. Folglich wäre es vermutlich ehrlicher, gar nicht erst von doppelblinder Begutachtung zu sprechen, da es falsche Erwartungen weckt.

Ein weiterer Nachteil der bisherigen Begutachtungspraxis ist, dass sie vermutlich tatsächlich im Durchschnitt schwache Beiträge aussortieren hilft, das aber zu Lasten der Veröffentlichung wirklich kreativer Beiträge geht. Der sogenannte α -Fehler, der auch als positiver Fehler oder Fehler erster Art bezeichnet wird, sinkt, während der β -Fehler bzw. negative Fehler oder Fehler zweiter Art steigt. Während der einzelne Gutachter kaum Anreize für eine Begutachtung in irgendeine systematische Richtung hat, sind die die Gutachter auswählenden Herausgeber der wissenschaftlichen Zeitschriften an deren Reputation interessiert. Da sich diese nur schwer messen lässt und in Deutschland z. B. keine systematische Zitationsanalyse mit der Berechnung von Impactfaktoren existiert, ist zumindest die Vermeidung von Kritik ein praktikabler Ersatz.

Seriöse Zeitschriften sollten nach Möglichkeit keinen Unsinn drucken, so dass im Zweifelsfall ein Beitrag lieber abgelehnt als angenommen wird, zumal eine hohe Ablehnungsquote als ein weiteres Signal für nicht unmittelbar beobachtbare Qualität interpretiert werden kann. Je renommierter eine Zeitschrift ist, desto mehr Gutachter werden in der Regel pro Beitrag eingesetzt. Doppelblinde Begutachtung setzt mindestens einen Gutachter voraus. Die besseren deutschen BWL-Zeitschriften senden jeden Beitrag an mindestens zwei Gutachter, bei den besten US-amerikanischen Journals können es vier oder sogar noch mehr sein. Die Entscheidung muss auch nicht „demokratisch“ fallen, etwa drei zu eins, sondern die Filterwirkung lässt sich noch erhöhen, wenn bereits ein ablehnendes Gutachten insgesamt zur Ablehnung führt. Willkürliche Annahmeentscheidungen einzelner Gutachter fallen dann nicht mehr ins Gewicht, wirklich schlechte und später erfolgreich kritisierbare Manuskripte werden kaum

⁷ Justice et al. (1998) zeigen, dass zumindest in ihrer Studie 32 % der anonymisierten Autoren erkannt werden und die Anonymisierung der Autoren für die Gutachten ohnehin keinen signifikanten Unterschied macht.

noch angenommen. Doch der α -Fehler lässt sich auf diese Weise nicht kostenlos reduzieren. Der Preis liegt in einem höheren β -Fehler, der Ablehnung guter bis exzellenter, vielleicht sogar nobelpreiswürdiger Beiträge. Gerade innovative und kreative Ideen sind riskant, da ihr Erfolg noch nicht feststeht. Entsprechende Beiträge werden also eher abgelehnt. Die besten Annahmehancen hat das gute Mittelmaß. Ein hoher Grad der Formalisierung, sei es mathematisch, sprachlich oder gedanklich, senkt das Risiko.⁸

Die Beiträge werden inhaltlich abgeschliffen, wozu eine weitere Eigenart des Begutachtungsprozesses beiträgt, nämlich die Auflagen der Gutachter. Jeder Gutachter kann beliebige Auflagen formulieren, von deren Erfüllung seine spätere Befürwortung der Annahme abhängig gemacht wird.⁹ Das Ergebnis ist, dass die veröffentlichten Beiträge nicht mehr den eingereichten entsprechen. Es handelt sich um eine Mischung der Ideen des Autors mit denen der Gutachter. Die Hauptarbeitslast trägt der Autor, der dann am Ende auch als einziger namentlich erwähnt wird, doch es sind nicht mehr allein seine Vorstellungen, die veröffentlicht werden, in manchen Fällen handelt es sich sogar um das genaue Gegenteil.¹⁰

4. Institutionelle Stärkung von Kritik

Das Dilemma wissenschaftlicher Kritik, dass ihre offene Äußerung wegen möglicher Vergeltungsmaßnahmen gefährlich ist, während für anonyme Kritik kein (Qualitäts-)Anreiz besteht, lässt sich im Zeitalter des Internets relativ einfach durch konsequentes Veröffentlichen und multiple bzw. zumindest doppelte Identitäten lösen. Eigene konstruktive Beiträge schreiben die Autoren weiterhin unter ihrem Namen, wodurch sie Reputation gewinnen oder gegebenenfalls bei schlechter Leistung auch verlieren können. Kritik, sei es in Form von Gutachten, Kommentaren oder auch Gegenartikeln, wird ebenfalls konsequent veröffentlicht, zumindest im Internet. Dies allein hat schon eine disziplinierende Wirkung auf die Kritiker, da eine schwache Argumentation öffentlich sichtbar würde und ihr Urheber sich dafür schämen müsste, selbst wenn sein Name nicht bekannt wird. Außerdem erlaubt dies das Äußern von Gegenkritik und initiiert möglicherweise eine öffentliche Debatte, wo bislang keine geführt wird.

⁸ Siehe Dilger (2012).

⁹ Dabei kommt es gelegentlich vor, dass selbst nach Erfüllung aller Auflagen, sogar vom Gutachter zugegeben, dieser trotzdem für die Ablehnung des Beitrags plädiert. Dies unterstreicht die Willkür des Prozesses, kommt aber zugleich nicht derart oft vor, dass die Erfüllung der Auflagen von vornherein keinen Sinn mehr machen würde.

¹⁰ Siehe Frey (2003) zur Kritik daran mit einem eigenen Lösungsvorschlag, der nicht so weit geht wie der hier im vierten Abschnitt entwickelte.

Als weitere Maßnahme sollte solche Kritik nicht völlig anonym erfolgen, sondern mit einem eindeutigen Pseudonym versehen werden. Mit dem Pseudonym kann ebenfalls Reputation gewonnen werden, nämlich als gekannter Kritiker. Dies allein würde schon einen Anreiz zu begründeter Kritik schaffen, der sich durch die Möglichkeit einer Verknüpfung der verschiedenen Identitäten noch deutlich steigern ließe. Dazu dürfte natürlich nicht der wahre Name mit dem Pseudonym öffentlich verknüpft werden, doch eine Bestätigung von Zeitschriften, dass jemand z. B. zu den besten zehn oder besten hundert der gerankten Gutachter bzw. Kritiker gehört, könnte etwa bei Berufungsverfahren Berücksichtigung finden.

Grundsätzlich sind Zeitschriften stärker für Kritik und Diskussionen zu öffnen. Nicht dass ein oder zwei anonyme Gutachter einen Beitrag für publikationswürdig befunden haben, zeugt von höchster wissenschaftlicher Qualität, sondern die Offenheit gegenüber jeder sachlichen Kritik, also potentiell den Einwänden aller Wissenschaftler. Nach diesem Kriterium kann eine herausgeberbegutachtete Zeitschrift besser sein als eine doppelblind referierte. Zeitschriften sind auch nicht per se besser als Bücher, können jedoch leichter die Erwiderung und Diskussion in einer Folgeausgabe ermöglichen. Schließlich könnten sich auf wissenschaftliche Kritik spezialisierte Zeitschriften (echte „Reviews“) oder auch Internetplattformen entwickeln.

Eine offene Publikationsplattform, die jeden Beitrag aufnimmt, dann aber der Kritik aussetzt und somit differenziert bewertet, ist außerdem deutlich effizienter als das bisherige Begutachtungssystem, wo viele Zeitschriften je für sich über die Annahme von Artikeln entscheiden. Wurde im dritten Abschnitt dargelegt, dass jede einzelne Zeitschrift bemüht ist, zu schlechte Beiträge durch den Begutachtungsprozess abzuwehren, was leider auch die Annahme besonders guter und innovativer Artikel hemmt, so gilt für die Gesamtheit aller Zeitschriften fast das Gegenteil. Auch ziemlich schlechte bis eindeutig fehlerhafte Artikel werden mit sehr hoher Wahrscheinlichkeit irgendwo angenommen, wenn der Autor nur beharrlich bleibt und nach jeder Ablehnung einen neuen Einreichungsversuch woanders unternimmt. Die z. T. hohen Ablehnungsquoten bei Topzeitschriften von weit über 90 % erscheinen in einem anderen Licht, wenn die Möglichkeit der wiederholten Einreichung berücksichtigt wird.

Die Gutachten sind gegenwärtig nicht nur anonym, sondern auch nicht öffentlich. Ein einmal abgelehntes Manuskript muss an anderer Stelle erneut von vorne geprüft werden, wenn beide Zeitschriften nicht zufällig auf den gleichen Gutachter zurückgreifen. Zumindest wenn Mängel relativ schwer zu entdecken sind, bleiben sie aus stochastischen Gründen irgendwann einmal unentdeckt. Der vervielfachte Begutachtungsaufwand erlaubt in jedem einzelnen Fall nur eine geringere Beurteilungsintensität. Bei einem öffentlichen Beurteilungsverfahren, wo

jeder Beitrag, alle Gutachten und Stellungnahmen im Internet veröffentlicht werden, vermeidet diese Ineffizienz. Während jetzt das bloße Erscheinen in einer renommierten Zeitschrift das wesentliche oder gar einzige Qualitätssignal darstellt, ist die Veröffentlichung selbst nach diesem Vorschlag kein Qualitätssignal mehr, während dafür die nachfolgende Diskussion eine sehr differenzierte und fundierte Qualitätseinschätzung erlaubt.

5. Fazit

Kritik ist wichtig für die Wissenschaft. Es muss differenziert werden, welche Ideen, Beiträge, Theorien und Wissenschaftler exzellent, welche gut, weniger gut oder einfach nur schlecht sind. Angesichts knapper Ressourcen fallen entsprechende Entscheidungen jeden Tag, doch nicht immer sachlich und explizit. Die offene und ehrliche Äußerung und Diskussion von Kritik wäre hier ein methodischer Fortschritt, der große inhaltliche Fortschritte nach sich ziehen dürfte. Um dies zu erreichen, müssen die Anreize zum sachlichen Kritisieren verbessert werden. Dazu wurden in diesem Beitrag einige Vorschläge entwickelt, doch viele weitere Ideen und vor allem deren versuchsweise, selbst für Kritik offene Umsetzung sind nötig. Es lohnt sich auch, von anderen Disziplinen zu lernen. Insbesondere die Naturwissenschaften sind in dieser Richtung oft schon weiter als die Wirtschafts- und Sozialwissenschaften.

Literatur

- Armbrüster, Thomas/Kieser, Alfred (2001): „Unternehmensberatung: Analyse einer Wachstumsbranche“, Die Betriebswirtschaft (DBW) 61, S. 688-709.
- Dilger, Alexander (2012): „Rigor, wissenschaftliche und praktische Relevanz“, Diskussionspapier des Instituts für Organisationsökonomik 3/2012, Münster.
- Frey, Bruno S. (2003): „Publishing as Prostitution? Choosing Between One's Own Ideas and Academic Success“, Public Choice 166, S. 205-223.
- Justice, Amy C./Cho, Mildred K./Winker, Margaret A./Berlin, Jesse A./Rennie, Drummond (1998): „Does Masking Author Identity Improve Peer Review Quality? A Randomized Controlled Trial“, JAMA 280, S. 240-242.
- Kant, Immanuel (1998): „Kritik der reinen Vernunft“, Hamburg (1. Auflage 1781, Riga).
- Kuhn, Thomas (1962): „The Structure of Scientific Revolutions“, Chicago; auf Deutsch: „Die Struktur wissenschaftlicher Revolutionen“, 21. Auflage 2001, Berlin.
- Lakatos, Imre (1977): „The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1“, Cambridge (UK).
- Mulligan, Adrian (2004): „Is Peer Review in Crisis?“, Perspectives in Publishing 2, S. 1-6.

Popper, Karl (1994): „Logik der Forschung“, 10. Auflage (1. Auflage 1934), Tübingen.

Spier, Ray (2002): „The History of the Peer-review Process“, Trends in Biotechnology 20, S. 357-358.

Diskussionspapiere des Instituts für Organisationsökonomik

Seit Institutsgründung im Oktober 2010 ist monatlich ein Diskussionspapier erschienen. Im Folgenden werden die letzten zwölf aufgeführt. Eine vollständige Liste mit Downloadmöglichkeit findet sich unter <http://www.wiwi.uni-muenster.de/io/de/forschen/diskussionspapiere.html>

- DP-IO 8/2017** Zur Stärkung wissenschaftlicher Kritik
Alexander Dilger
August 2017
- DP-IO 7/2017** Theoretische Erklärungsansätze für die Entsprechenserklärungen zu Abfindungen für Vorstandsmitglieder
Ute Schottmüller-Einwag
Juli 2017
- DP-IO 6/2017** Doping in Teams
A Simple Decision Theoretic Model
Alexander Dilger
Juni 2017
- DP-IO 5/2017** Wirtschaftsethische Überlegungen zur Finanz- und Eurokrise
Alexander Dilger
Mai 2017
- DP-IO 4/2017** Vor- und Nachteile verschiedener Arten von Drittmitteln
Alexander Dilger
April 2017
- DP-IO 3/2017** The Euro from a Business Perspective
Alexander Dilger
März 2017
- DP-IO 2/2017** Is Trustworthiness Written on the Face?
Alexander Dilger/Julia Müller/Michael Müller
Februar 2017
- DP-IO 1/2017** Physical Constitution Matters for Athletic Performance and Salary of NBA Players
Linn-Brit Bakkenbüll
Januar 2017
- DP-IO 12/2016** Übereinstimmung von Entsprechenserklärungen und Erklärungsentsprechen bei Abfindungen
Ute Schottmüller-Einwag
Dezember 2016
- DP-IO 11/2016** Willingness to Pay and Accept for Hosting Olympic Games in Germany
Linn-Brit Bakkenbüll/Alexander Dilger
November 2016
- DP-IO 10/2016** 6. Jahresbericht des Instituts für Organisationsökonomik
Linn-Brit Bakkenbüll/Alexander Dilger
Oktober 2016
- DP-IO 9/2016** Herausforderungen der Flüchtlingskrise für Hochschulen
Alexander Dilger
September 2016



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

