



Einsichten eines Wissenschaftsnarren (35)

Back to the Future: Von industrieller zu inhaltlicher Forschungsbewertung

Wie effizient kann ein Bewertungssystem sein, dessen Messgrößen sich von den Inhalten, der Relevanz und der Qualität der Forschung verabschiedet haben? Sie haben es gemerkt: Wir sprechen von der schrägen Forschungsevaluation anhand reiner Zahlen aus bibliometrischen Daten und Drittmiteleinwerbung. Ist solch ein System, das sich weltweit durchgesetzt hat, überhaupt noch reformierbar? Unser Wissenschaftsnarr meint: Aber sicher!

Wissenschaft verschlingt massiv gesellschaftliche Ressourcen, nicht nur finanzielle. Insbesondere für die akademische Forschung, die sich selbst verwaltet und sich gerne auf die im Grundgesetz verankerte Forschungsfreiheit beruft, stellt sich damit die Frage, wie sie die Mittel einteilt, die ihr von der Gesellschaft zur Verfügung gestellt werden. Es gibt keine natürliche Begrenzung, wie viel geforscht werden könnte – aber sehr wohl eine Beschränkung der Mittel, die die Gesellschaft für Forschung einsetzen kann und will. Welche Forschung soll also gefördert, welche Wissenschaftler ins Brot gesetzt werden?

Für die Beantwortung dieser zentralen Fragen für den akademischen Betrieb haben sich über viele Jahrzehnte gewisse Mechanismen evolutionär entwickelt. Diese Mechanismen steuern aber



Foto: BIH/Thomas Rafalzyk

Ulrich Dirnagl

leitet die Experimentelle Neurologie an der Berliner Charité und ist Gründungsdirektor des QUEST Center for Transforming Biomedical Research am Berlin Institute of Health. Für seine Kolumne schlüpft er in die Rolle eines „Wissenschaftsnarren“ – um mit Lust und Laune dem Forschungsbetrieb so manche Nase zu drehen.

nicht nur die Verteilung der Mittel in und zwischen den Institutionen, sondern letztendlich auch die Inhalte und die Qualität der Forschung. Den individuellen Wissenschaftlern, die in der Academia nicht nur ihrem Forscherdrang nachgehen, sondern auch ihren Lebensunterhalt verdienen, geht dies in Fleisch und Blut über – sie halten das für so etwas wie eine natürliche Ordnung. Die Verteilungs- und Leistungsbewertungsmechanismen und die dazugehörige Indikatorik bestimmen ihren Tagesablauf samt der Art und Weise, wie sie forschen, mehr als die tägliche Lektüre von Fachliteratur, der Blick durchs Mikroskop oder Kongressvorträge. Auch wenn das den Wenigsten wirklich bewusst ist.

In der letzten Folge (LJ 12/2020: 18-19) hat der Wissenschaftsnarr die Frage gestellt, wie sich das heute weltweit etablierte Karriere-, Belohnungs- und Begutachtungssystem entwickeln konnte – von den Anfängen der modernen Wissenschaft im 17. Jahrhundert

»Wie konnte es kommen, dass das Motto zur Beurteilung von Forschern lautet: „Sag mir deinen JIF – und ich sage Dir, ob Du ge- oder befördert wirst.“«

bis heute. Wie es dazu kommen konnte, dass quantitative Indikatoren wie der *Journal-Impact-Factor* (JIF) und die Höhe der Drittmiteleinwerbung bei der Beurteilung von Forschern und deren Anträgen eine wichtigere Rolle spielen als die Inhalte, Relevanz oder Qualität der Forschung selbst. Frei nach dem Motto: „Sag mir deinen JIF – und ich sage Dir, ob Du ge- oder befördert wirst.“

Dabei stellte sich heraus, dass dieses Beurteilungssystem nur wenige Dekaden alt ist; vermutlich ist erst eine Generation von Wissenschaftlern komplett in ihm sozialisiert worden. Das System erwuchs im Wesentlichen aus zwei Entwicklungen. Zum einen aus der Industrialisierung und massiven Ausweitung von akademischer Forschung. Diese wiederum ist das Resultat ihres eigenen immensen Erfolgs, aber gleichzeitig auch der mit diesem Erfolg abnehmenden Effizienz von Forschung geschuldet. Denn weil die „Früchte der Erkenntnis“ immer höher hängen, benötigt es einen immer größeren Einsatz an Forschung, um den Erkenntnisgewinn pro eingesetztem Förderbetrag konstant zu halten, wenn nicht zu steigern. Zusammengefasst führt dies zu einer immer weniger beurteilbaren Flut von Forschern, Projekten, Anträgen und Artikeln. Um da noch durchzukommen, benötigen wir einfach und schnell zu erhebende Bewertungskriterien. Am besten welche, die man anwenden kann, ohne sich die Mühe zu machen, die eigentlichen Inhalte der Wissenschaft zu beurteilen.

Die zweite wesentliche Triebfeder der Entstehung des heutigen Beurteilungssystems ist der nachvollziehbare Wunsch nach Verteilungsgerechtigkeit. Wir wünschen uns objektive Kriterien, die repro-

Sämtliche Folgen der „Einsichten eines Wissenschaftsnarren“ gibt es unter www.laborjournal.de/rubric/narr

duzierbar sind und nicht der Willkür unterworfen – und die eine eindeutige Diskriminierung zwischen Bewerbern oder Anträgen erlauben, im besten Fall sogar ein einfaches Ranking. Niemand soll gefördert werden, weil sich ein mächtiger Mentor hinter den Kulissen eingemischt hat. Sondern weil man etwas geleistet hat, das jeder nachvollziehen kann.

Und schon sind wir beim JIF und den akkumulierten Drittmitteln. Einfach, objektiv, quantifizierbar, nachvollziehbar. Man muss weder einen Artikel eines Kandidaten gelesen haben noch dessen ganzen Lebenslauf. Ein Blick auf die Literaturliste (für die Doo-fen gerne mit dem auf 2 oder 3 Nachkommastellen genauen JIF hinter jedem Journal-Namen), dazu noch auf die Aufreihung der Drittmittel – das reicht völlig. Wer Übung hat und häufiger in Kommissionen sitzt oder begutachtet, schafft das locker in drei Minuten pro Kandidat.

»Forschung kann nur beurteilen, wer sich kompetent und konkret mit deren Inhalten, Methoden, Ergebnissen und Interpretationen auseinandersetzt.«

Natürlich liegt genau hier der Hase im Pfeffer: Wie effizient und nützlich kann ein Verteilungs- und Bewertungssystem sein, dessen Mess- und Steuergrößen sich von den Inhalten, der Relevanz und der Qualität der zu steuernden Forschung verabschiedet haben? Nicht nur der Wissenschaftsnarr schlägt hier Alarm. Die Spatzen pfeifen es von den Dächern, dass es so nicht weitergehen kann.

Wie aber könnte man es besser machen? Ist ein System überhaupt reformierbar, das sich weltweit durchgesetzt hat und dazu doch offensichtlich funktioniert – Stichwort CRISPR, SARS-CoV2-Vakzine *et cetera*? Nur ein Narr kann sich erlauben, diese Frage mit einem klaren „Ja“ zu beantworten – und gleich noch ein paar praktische Vorschläge hierfür zu formulieren.

Vorweg zunächst drei Prämissen:

1. Forschung kann nur beurteilen, wer sich kompetent und konkret mit deren Inhalten, Methoden, Ergebnissen und Interpretationen auseinandersetzt. Das ist natürlich sehr unangenehm, denn das ist aufwendig, kann nicht automatisiert werden und ist nicht quantifizierbar.

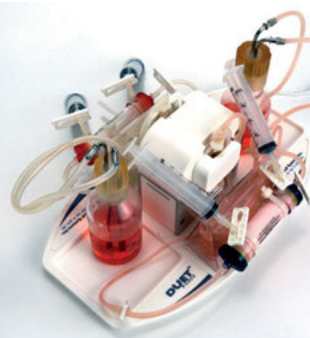
2. Wir dürfen den gesellschaftlichen Einsatz an Ressourcen für Forschung nicht reduzieren, sondern müssen die vorhandenen Ressourcen effizienter einsetzen. Man könnte nämlich darauf kommen, einfach weniger Forschung zu fördern. Auf diese Weise könnte man den Output so weit reduzieren, bis dieser am Ende wieder inhaltlich bewertbar würde. Allerdings würde man damit die Uhr mehr als hundert Jahre zurückstellen und eine wissenschaftliche Eiszeit induzieren – das geht natürlich gar nicht!

3. Die nötigen Veränderungen im Bewertungs- und Verteilungssystem müssen von außen kommen, also von den staatlichen Förderinstitutionen, den Hochschulen und den außeruniversitären Wissenschaftsorganisationen. Die Wissenschaftler, die ihren Weg in die Academia suchen, haben schließlich keine andere Wahl, als sich den Bedingungen der Konkurrenz um Fördermittel und berufliche Positionen zu stellen. Sie sind ja das Objekt der Bewertungsmechanismen.

Womit könnte man beginnen? Um eine inhaltliche und qualitative Bewertung von Forschungsleistungen durchzusetzen, müsste man zunächst die Verwendung von abstrakten Indikatoren (JIF, Drittmittelinwerbung *et cetera*) gezielt verhindern – und nicht nur deren sparsamen und nur unterstützenden Gebrauch anmahnen! Dies be-

Hollow Fiber Bioreactor

Scale-up productions from mammalian cells using minimum incubator space



Monoclonal Antibodies
Extracellular Vesicles
Recombinant Proteins

The FiberCell Hollow Fiber Bioreactor can continuously grow practically any kind of cell at very high density: hybridomas, CHO, HEK, BHK, S2, Huh.7, THP1, A549 and MSCs...

10⁹ cells or more cultured continuously

Only requires one shelf in a standard incubator

Uses simplified chemically-defined, protein-free media



KDBIO S.A.S., 6 r. Iris, 67370 Berstett
www.kdbio.com, info@kdbio.com, 03.88.26.12.86
European distributor for FiberCell Systems Inc.



deutet konkret: Die Angabe von JIF, h-Faktor und Co. verbieten und durch die obligatorische Verwendung von Narrativen zur Beschreibung des eigenen Beitrages ersetzen. Das Zitieren eigener Arbeiten in Lebensläufen und Anträgen sollte nur unter Angabe von Titel, Autoren sowie eines *Identifiers* wie etwa der PMID (*PubMed Identifier*) ohne Nennung des jeweiligen Journal-Namens erfolgen. Damit kann die Referenz aufgerufen und gelesen werden, das pure Durchsuchen von Literaturlisten nach Journal-Namen ist dann aber nicht mehr möglich. Diese Kurznarrative würden auch ganz natürlich zu einer Beschränkung auf wenige relevante Literaturstellen führen. Denn wer wollte schon mehr als zehn oder mehr davon schreiben?

Eine Fokussierung auf Erst- und Letztautor-Positionen ist dann auch nicht mehr nötig und sollte ganz entfallen. Schließlich handelt es sich hierbei ohnehin um eine potenziell schädliche Fiktion: Heutzutage liefern zu jeder relevanten biomedizinischen Arbeit eine Vielzahl von Wissenschaftlern verschiedenartigste Beiträge. Diese lassen sich nicht auf zwei Positionen in der Autorenleiste reduzieren, die noch dazu gar nicht eindeutig definiert sind. Die geteilte Autorschaft mit Sternchen ist nichts weiter als der alberne Versuch, sich um diese Erkenntnis herumzumogeln.

Damit zusammenhängend müssten auch die Mindestanzahlen von Publikationen fallen, wie sie derzeit für Promotion, Habilitation *et cetera* gefordert werden. All dies provoziert nur das *Slicing* von Studien in kleinere Einheiten, die Inflation von Publikationen, die keiner braucht, unsinnige und unnötige Diskussionen um Autorenpositionen – sowie noch andere Unsitten. Stattdessen sollte ein Narrativ den wissenschaftlichen Beitrag individueller Wissenschaftler darlegen. Ob dieser dann Promotions- oder Habilitations-würdig ist, müssen die zuständigen Kommissionen in einer inhaltlichen Auseinandersetzung mit dem *Ceuvre* der Kandidaten entscheiden, aber nicht wie derzeit üblich aus dem Studium eines *Spreadsheet-Rankings* ableiten.

Bei dieser Gelegenheit sollte man dann gleich versuchen, zu alphabetischen Autorenlisten überzugehen, wie in den Multiautor-Kollaborationen der Hochenergie-Physik schon lange erfolgreich praktiziert. Dafür gibt es unter <https://casrai.org/credit> bereits eine hervorragende Taxonomie, die sich auch für die Lebenswissenschaften eignet.

»Die Karrierewege in der Academia müssen sich ändern. 83 Prozent des wissenschaftlichen Personals sitzt auf befristeten Stellen!«

Letztendlich ergeben sich Reputation und Renommee von Wissenschaftlern doch aus ihren inhaltlichen Beiträgen und ihrem *Standing* in der Community. Daher sollten auch *Reviews* und Beurteilungen durch *Peers* Berücksichtigung finden, die nach deren Publikationen entweder bei den Journalen (etwa nach *Post-Publication-Review*) oder aber auch auf sozialen Medien publiziert werden. *Science Twitter* ist in vielen Feldern bereits heute wesentlich transparenter, nachvollziehbarer, aktueller und damit auch *wissenschaftlicher* als althergebrachte Formate des Diskurses wie etwa *Letters to the Editor* oder Ähnliches.

Die Qualitätskontrolle wissenschaftlicher Publikationen findet ohnehin in den sozialen Medien effizienter statt als im klassischen *Peer Review*. Ein Beispiel hierfür ist, dass die Qualitätsprobleme, die am Ende zur Retraktion von Papers aus hochrangigen Journalen führen, bereits seit einiger Zeit zuallererst auf Twitter oder in Blogs exponiert werden. Und vorher im *Peer Review* regelhaft übersehen wurden.

Die oben genannten Maßnahmen würden bereits eine massive Reduktion der Artikelflut bewirken, wodurch eine Auseinandersetzung mit deren Inhalten erleichtert wird. Inhalte und Qualität können dann auch tatsächlich wissenschaftliche Reputation und Renommee bestimmen, und nicht Proxies wie JIF und Drittmittel.

Es fehlt aber noch etwas Wesentliches: Die Karrierewege in der Academia müssen sich ändern. 83 Prozent des wissenschaftlichen Personals sitzt auf befristeten Stellen! Der immense Konkurrenzdruck, überhaupt im System bleiben zu können oder es gar von der Basis der Pyramide an die Spitze zu schaffen, führt zur Selektion von Eigenschaften, die weder förderlich für Qualität noch für Kooperation in der Wissenschaft sind. Die Pyramide muss vielmehr zu einem Trapez geformt werden. Dabei muss die Spitze flacher, und die Basis etwas weniger breit werden. Das bedeutet aber auch, weniger PhD-Studenten (als „billige“ Arbeitskräfte) ins System einzuschleusen als bisher. Wer den keineswegs unbeschwerlichen Weg in die akademische Berufswelt nimmt, muss die reelle Chance haben, durch gute Wissenschaft (und nicht nur durch drei „Top-Publikationen“) langfristig ein Auskommen zu haben.

»Da Innovation nicht vorhersagbar ist und Begutachtung den Mainstream begünstigt, sollte ein Teil der Förderung in Lotterien vergeben werden.«

Und nun setzt der Narr zum letzten Schlag an: Nach Einführen eines rein inhalts- und qualitätsorientierten Bewertungssystems sowie Kappen der akademischen Karriere-Pyramide fehlt als drittes Element noch... der Zufall! Da echte Innovation nicht vorhersagbar ist und jeder wie auch immer geartete Begutachtungsprozess tendenziell den Mainstream begünstigt, sollte ein Teil der Förderung in Lotterien vergeben werden. Ja, Sie haben richtig gelesen: Verlost werden! Wer Näheres dazu wissen will, dem sei ein früherer Beitrag des Wissenschaftsnarren empfohlen (*LJ* 04-2019: 18-19). Solch ein Vergabemodus würde uns auch mehr Zeit zum Forschen lassen, weil ein Teil der Antragschreiberei und deren Begutachtung wegfallen würde. Sicher würde vieles gefördert werden, das eher mittelmäßig ist und nicht den versprochenen Durchbruch bringt – doch das ist ja jetzt auch schon so. Die Wahrscheinlichkeit aber, dass tatsächlich mal etwas bahnbrechend Neues gefördert würde, stiege indes erheblich.

Doch wie realistisch sind solch närrische Gedankenspiele? Kaum zu glauben, aber die DFG arbeitet derzeit tatsächlich an einem Positionspapier, das – abgesehen von der Lotterie – all dies und sogar noch mehr umzusetzen empfehlen soll, was der Narr sich hier gerade eben zusammengespinnen hat. Und das ist nicht nur eine der üblichen „Denkschriften“, mit denen man zeigt, dass man ein ‚Problem‘ erkannt hat und man sich viel vornimmt – aber wenig tut, weil alles so furchtbar komplex ist. Die DFG empfiehlt sich darin selbst, gleich und ganz konkret mit der Umsetzung zu beginnen! Wenn dieses Papier verabschiedet wird, reiht sich also endlich auch unser wichtigster Forschungsförderer ein in die weltweite Riege der Fördergeber und Institutionen – wie etwa der Wellcome Trust oder die holländische ZonMw –, die es ernst meinen damit, dass es so nicht weitergehen kann.

Sogar die Lotterie wird übrigens mancherorts schon ausprobiert, zum Beispiel bei der VolkswagenStiftung (*siehe S. 72-75 in diesem Heft*). Vielleicht ist also gar keine Zeitmaschine mehr nötig, und die Zukunft hat schon (ein bisschen) begonnen?

Weiterführende Literatur und Links finden sich wie immer unter: <http://dirnagl.com/lj>.