

Akzeptierte Manuskriptfassung (nach peer review) des folgenden Artikels:

[Abele-Brehm, A. & Bühner, M. \(2016\). Überlegungen zur Optimierung von Berufungsverfahren in der Psychologie. Psychologische Rundschau, 67 \(4\). doi:10.1026/0033-3042/a000333](#)

© Hogrefe Verlag, Göttingen 2016

Diese Artikelfassung entspricht nicht vollständig dem in der Zeitschrift veröffentlichten Artikel. Dies ist nicht die Originalversion des Artikels und kann daher nicht zur Zitierung herangezogen werden.

Die akzeptierte Manuskriptfassung unterliegt der Creative Commons License CC-BY-NC.

# Überlegungen zur Optimierung von Berufungsverfahren in der Psychologie

Andrea Abele-Brehm & Markus Bühner<sup>1</sup>

Friedrich-Alexander Universität Erlangen-Nürnberg

Ludwig-Maximilian Universität München

Die Qualität unserer Wissenschaft wird maßgeblich über die Personen, die psychologische Professuren besetzen, bestimmt. Gute Besetzungen können das gesamte Institut bzw. Fakultät voranbringen. Fehlbesetzungen können jahrzehntelange negative Auswirkungen auf ein gesamtes Institut bzw. eine Fakultät haben. Die große Relevanz der Aufgabe „Auswahl der geeignetsten Person für die Professur“ korrespondiert häufig jedoch nicht der Professionalität der Berufungsverfahren. Der vorliegende Beitrag möchte die Diskussion über die Validität von Entscheidungen in Berufungsverfahren initiieren und einige Überlegungen präsentieren, wie diese unter Berücksichtigung vielfältiger Anforderungen an Professuren gesteigert werden könnte.

Idealerweise sollte ein/eine neu zu berufende Kollege/Kollegin exzellent in der Forschung sein, in der Lehre die Studierenden inspirieren und begeistern, Forschungsprojekte sowohl einwerben als auch managen, Führungs- und soziale Kompetenz allgemein besitzen und obendrein sympathisch und engagiert sein. Ist das so? Häufig spielen solche umfassenden Überlegungen kaum eine Rolle, man konzentriert sich auf die Forschungsleistung und hofft, dass der Rest schon passen wird.

Entsprechend sehen auch Berufungsverfahren aus: Eine heterogen zusammengesetzte und meist viel zu große Kommission wählt anhand einer von einem Kommissionsmitglied erarbeiteten Synopse auf Basis hauptsächlich quantitativer Indikatoren (z. B. Anzahl an Publikationen; Drittmittelvolumen; h-Index) in einer Sitzung aus den (30 bis 60) Bewerbungen die (vier bis acht) Bewerber/innen aus, die zu einem Vorstellungsgespräch eingeladen werden. Von diesen Personen fordert die Kommission häufig, aber nicht immer einige Publikationen an, die von jeweils zwei Personen angesehen werden. Die Kommission hört sich den Bewerbungsvortrag (Forschung, manchmal auch zusätzlich Lehre) an und diskutiert ein wenig öffentlich darüber. Anschließend unterhält sie sich etwa 30 Minuten mit dem/der Bewerber/in. Später entscheidet sie, welche drei (oder vier) Personen extern begutachtet werden sollen und bestimmt nach Eingang der Gutachten<sup>2</sup> die Reihung auf der Berufsungsliste.

Im Vergleich zu den wünschenswerten Merkmalen und auch im Vergleich zu der großen Bedeutung, die die Berufung eines/einer Kollegen/Kollegin für ein Institut hat ist das Auswahlverfahren mager, die Auswahlkriterien sind eng, wichtige Bereiche werden nicht erkundet. Schlimmer noch, mit bestimmten Auswahlkriterien werden Entwicklungen in unserem Fach befördert, die die Qualität der Forschung und den Stellenwert der Psychologie als Wissenschaft beeinträchtigen können.

Unsere Thesen:

---

<sup>1</sup> Korrespondenz an [andrea.abele-brehm@fau.de](mailto:andrea.abele-brehm@fau.de) oder [buehner@psy.lmu.de](mailto:buehner@psy.lmu.de) Wir danken Mario Gollwitzer, Conny Antoni und Bianca Vaterrodt für hilfreiche Hinweise zu diesem Papier.

<sup>2</sup> Die Gutachter werden meist aufgefordert, Forschungsleistung, Lehrbefähigung und Drittmittelinwerbung vergleichend zu bewerten.

1. Bei universitären Personalauswahlverfahren haben rein quantitative Produktivitätsindikatoren ein zu hohes Gewicht. Das ist unzureichend und tendenziell sogar schädlich. Die Qualität der Kandidatinnen und Kandidaten in Forschung, Lehre und Selbstverwaltung muss nach multiplen Kriterien beurteilt werden.
2. Bei Berufungsverfahren müssen überfachliche Kompetenzen wie soziale Kompetenzen (z.B. Kommunikation, Führung, Kooperation) und Managementkompetenzen (z.B. organisations- und strategische Kompetenzen) mehr berücksichtigt werden.
3. Die Entscheidungsfindung in Berufungsverfahren kann verbessert werden, wenn diagnostische Methoden jenseits von Vortrag und – unstrukturiertem – Gespräch in der Berufungskommission zum Einsatz kommen.

### **Warum sind quantitative Produktivitätsindikatoren keine hinreichenden Kriterien für Forschungsexzellenz und welche Alternativen gibt es?**

In Berufungsverfahren werden meist quantitative Indizes herangezogen, um die Forschungsexzellenz eines/einer Wissenschaftlers/Wissenschaftlerin zu bewerten. Es werden die Anzahl der Publikationen, der h-Index, die Zitationsrate, die Anzahl von Publikationen in den sog. Top Journals des Feldes und/oder direkt der Impact Faktor des Publikationsorgans herangezogen. Alle diese Indikatoren erlauben zwar auf den ersten Blick einen einfachen Vergleich zwischen Bewerber/innen, sie haben aber auch deutliche Nachteile wie schon mehrfach aufgezeigt wurde (vgl. Schönbrodt et al., 2015)<sup>3</sup>.

Die starke Berücksichtigung solcher quantitativer Indikatoren hat darüber hinaus unerwünschte Konsequenzen auf das Publikationsverhalten und letztendlich auch auf die Forschungsqualität. So ist der Druck auf Jungwissenschaftler/innen, möglichst viel zu publizieren („publication bias“; „p-hacking“; vgl. Simmons, Nelson & Simonsohn, 2011; Ulrich et al., 2016, im Druck), extrem gestiegen und man erscheint erst ab einer bestimmten Anzahl an Publikationen konkurrenzfähig; so bewerten Zeitschrifteneditoren die Passung einer Einreichung zum jeweiligen Journal u.a. danach, wie „sexy“ bzw. „unerwartet“ ein Resultat ist, da entsprechende Artikel mehr Aufmerksamkeit erhalten; und so gibt es immer mehr Zeitschriften, die immer mehr Gutachter/innen für immer mehr Einsendungen erfordern und die von immer weniger Personen gelesen werden, um nur einige Beispiele zu nennen. Man kann es zwar nicht beweisen, aber es liegt doch nahe, dass die „Replikationskrise“ nicht unabhängig von dieser Art der Vermessung von Forschungsexzellenz ist.

Eine Alternative zum gängigen Vorgehen wäre, die wichtigsten drei Publikationen der letzten fünf Jahre (ausgewählt durch Bewerber/in) sorgfältig zu lesen und in der Berufungskommission zu diskutieren. Dies bedeutet einen höheren Aufwand, der sich aber aus unserer Sicht lohnt. Im Nachgang können bei vergleichbarer Qualität auch Produktivitätskriterien herangezogen werden. Beispielsweise kann eine Kombination aus Anzahl der Publikationen nach der Promotion pro Jahr, mittlere Zitationsrate pro Jahr und kumulierter Impact als nachgelagertes Entscheidungskriterium verwendet

---

<sup>3</sup> Der Impact-Faktor ist z.B. leicht manipulierbar durch selektives Annehmen von Beiträgen in sog. Top Zeitschriften. Der h-Index wächst mit dem Alter, benachteiligt Jungwissenschaftler und verlangt eine ausgewogene Publikationsstrategie, bei der einzelne herausragende Artikel sogar hinderlich sein können. Die Anzahl der Publikationen variiert mit der Publikationsstrategie einer Arbeitsgruppe, z.B. inwiefern bei jedem Artikel alle Mitglieder der Arbeitsgruppe auf die Autorenliste kommen. Auch Universitätsleitungen schenken den genannten Indikatoren zu viel Aufmerksamkeit. Als kleine Anekdote: Peter Higgs, der Nobelpreisträger für Physik aus dem Jahre 2013, vermutet, dass er heute keinen akademischen Beruf mehr bekommen würde, weil er quantitativ zu wenig produktiv war.

werden. Auch Aspekte von transparenter Forschung (z.B. Offenlegung publizierter Daten in Repositorien) könnten berücksichtigt werden. Als Zeitraum könnten die letzten fünf Jahre herangezogen werden bei entsprechender Berücksichtigung von Erziehungszeiten für Kinder. Würde sich aus der qualitativen Bewertung der drei besten Arbeiten der letzten fünf Jahre ein Gleichstand zwischen zwei Bewerberinnen oder Bewerbern ergeben, dann könnten drei Rangreihen zur weiteren Differenzierung gebildet und gemittelt werden: (1) durchschnittliche Anzahl der Publikationen pro Jahr nach der Promotion, (2) mittlere Zitationsrate, (3) mittlerer Impact der Zeitschriften. Ein aus unserer Sicht noch besseres Kriterium wäre, insbesondere bei einem Bewerberfeld aus stark unterschiedlichen Bereichen (z.B. Persönlichkeitspsychologie und Diagnostik mit einem biologischen Schwerpunkt), die Anzahl an Top-Publikationen im Feld in den letzten fünf Jahren zu betrachten. Somit wird eine Feldbereinigung vorgenommen, die berücksichtigt, dass beispielsweise Psychological Assessment und Brain stark unterschiedliche Impact Faktoren haben, obwohl beide Zeitschriften im Feld jeweils sehr angesehen sind.

Auch die Summe und Art eingeworbener Drittmittel spielt in Berufungsverfahren eine wichtige Rolle – und ist für Hochschulleitungen von besonderer Bedeutung. Wir halten den Faktor der eingeworbenen Drittmittel ebenfalls für wichtig. Allerdings sollte die Summe der bisher eingeworbenen Drittmittel nicht absolut, sondern relativiert auf den Durchschnittswert im Fach betrachtet werden. Auch sollte neben der Summe auch der „Output“ der Projekte betrachtet werden, z.B. die aus dem Projekt entstandenen Veröffentlichungen oder bei Forschung z.B. im Industriekontext, welcher soziale oder technische Impact bzw. Impact für die Lehre (z.B. durch Praxisprojekte) mit dem Drittmittelprojekt verbunden ist. Die bloße Summe ist nur ein bedingt guter Indikator für die Forschungsexzellenz.

### **Wie erfasst man Exzellenz in der Lehre?**

Während die universitäre Lehre noch vor wenigen Jahren in Berufungsverfahren kaum thematisiert wurde, nimmt sie heute einen höheren Stellenwert ein und das Einreichen von Lehrevaluationen bei Bewerbungsverfahren ist Standard geworden. Aber Lehrevaluationen durch die Studierenden sind mit vielen Fehlern behaftet (vgl. Wolbring, 2013).<sup>4</sup>

Eine aus unserer Sicht günstigere oder ergänzende Messung könnte dadurch erzielt werden, dass man von allen Bewerberinnen und Bewerbern einen genau spezifizierten Lehrvortrag zum gleichen Thema halten lässt, bei dem das Kriterium, d.h. was die Studierenden nach diesem Lehrvortrag gelernt haben sollten, klar definiert ist. Die Fachlichkeit des Lehrvortrags sollte von Experten und dessen Verständlichkeit von Studierenden bewertet werden. Auch eine Liste von Fragen zum Lehrvortrag, die allen Bewerberinnen und Bewerbern gestellt werden, kann hilfreich sein.

Wir halten auch die Erhebung eines Lehrkonzepts im Sinne hochschuldidaktischer Kompetenzen (Arten von Lehrveranstaltungen; Arten der studentischen Beteiligung, etc.) für sinnvoll. Schließlich gibt es an vielen Universitäten sowie vom Hochschullehrerverband Angebote zum Training hochschuldidaktischer Kompetenzen und Bewerber/Bewerberinnen, die entsprechende Erfahrungen mitbringen,

---

<sup>4</sup> Z.B. ist die Note der Hochschulzugangsberechtigung der Studierenden ist ein guter Prädiktor für das spätere studentische Urteil; korrelieren studentische Urteile zur Qualität von Einführungsvorlesungen statistisch überzufällig mit den Bewertungen von Lehrveranstaltungen durch dieselbe Person in höheren Semestern; werden Wahlveranstaltungen besser evaluiert als Pflichtveranstaltungen; ist das Vorinteresse am Veranstaltungsthema für eine positive Bewertung entscheidend. Ergänzende Kritikpunkte sind, dass Lehrevaluationen durch die Studierenden von denjenigen vorgenommen werden, die die Veranstaltung regelmäßig besuchen (also eine spezifische Selektion), dass Lehrevaluationen mit der empfundenen Schwierigkeit der Leistungsprüfung korrelieren und dass einer der wichtigsten Prädiktoren für die Bewertung der Lehre die Sympathie der Studierenden für die jeweilige Lehrperson ist.

bzw. entsprechende Zertifikate vorweisen können, könnten im Bereich Lehre ein „prä“ erhalten. Dies hätte neben der direkten Relevanz für die Qualität der Lehre auch Signalwirkung dahingehend, dass die didaktische Ausbildung einen ebenso hohen Stellenwert hat wie Kurse in Mplus oder R oder zu Strukturgleichungsmodellen.

### **Warum sollten Kriterien wie soziale Kompetenzen und Managementkompetenzen bei Berufsentscheidungen eine Rolle spielen?**

Erfolgreiche Forschung ist heute meist kooperative Forschung und hat mit der Humboldt'schen Prämisse der „Forschung in Einsamkeit und Freiheit“ nur noch wenig zu tun. Es geht um Zusammenarbeit mit den eigenen Mitarbeitern und Mitarbeiterinnen in einem mehr oder weniger großen Team, mit Kollegen und Kolleginnen innerhalb und außerhalb des eigenen Instituts, intra- und interdisziplinär, national und international. Forschungsverbände werden immer wichtiger, insbesondere größere und gut vernetzte Einheiten können gleichzeitig qualitativ hochwertig und mit der nötigen Power forschen. Aber: Forschungsverbände erfordern strategische Überlegungen zur Ausrichtung, sie erfordern die Bereitschaft und die Fähigkeit zu Kooperation. „Soft Skills“ wie kommunikative, Führungs- sowie Managementkompetenzen allgemein sind gefragt.<sup>5</sup>

Ähnliches gilt für die Lehre. Bachelor- und Masterprogramme, auch Graduiertenprogramme müssen unter den Beteiligten abgestimmt werden, Lehrinhalte sind zu koordinieren. Wenn sich Institute in Masterstudiengängen oder in Graduiertenprogrammen spezialisieren, können Kooperationen über Universitäten hinweg sinnvoll werden, d.h. „Lehrverbände“. Auch hierfür bedarf es strategischer, Management-, Führungs- und insbesondere Kommunikationskompetenzen.<sup>6</sup>

Betrachten wir als Beispiel Führungskompetenzen. Natürlich gibt es keinen optimalen Führungsstil, kompetente Führung ist auf die Anforderungen für die jeweilige Position bezogen. Führungsaufgaben von Professorinnen und Professoren erfordern insbesondere soziale Kompetenzen bei Personalführung, bei Teamführung, beim Konfliktmanagement und beim Umgang mit Kritik. Bei der Personalführung sind inspirierende Gespräche, fachliche Anregungen, Unterstützung und persönliche Wertschätzung hilfreich, wie sie in der Führungsliteratur unter dem Stichwort transformationale Führung beschrieben werden (vgl. Wegge & Rosenstiel, 2014). Je nach Person und Situation können jedoch auch konkrete Zielvereinbarungen und spezifisches Feedback nötig sein, so dass die Führungskraft die Fähigkeit mitbringen muss, die jeweiligen Anforderungen einschätzen und darauf flexibel eingehen zu können. Da mit einer Professur auch Kooperationsaufgaben verbunden sind, ergeben sich in dieser Position auch die Anforderungen, in Teams konstruktiv zusammenarbeiten und im Sinne „geteilter Führung“ dazu beitragen zu können, dass das jeweilige Team seine Ziele erreicht.

---

<sup>5</sup> Der Bericht zur Qualitätssicherung (Fiedler et al., 2016) enthält den Vorschlag, einen Preis für gelungene Forschungsoperation einzurichten. Wir halten diese Idee für sehr sinnvoll.

<sup>6</sup> Kommunikative und Management-Kompetenzen sind auch im Bereich fachpolitischer Aktivitäten wichtig. In dem Maße, in dem sich die Psychologie professionalisiert und als forschungs- und teilweise auch gesellschaftspolitischer „Player“ aktiv wird, braucht es Personen, die über die genannten Fertigkeiten verfügen. Managementkompetenzen und kommunikative Fähigkeiten im politischen Kontext können ein Fach extrem voranbringen (vgl. z.B. die laufende Neuregelung des Psychotherapeutengesetzes oder die Neuregelung fachgutachterlicher Tätigkeiten vor Gericht). Wir sind als Fach darauf angewiesen, kompetent vertreten zu werden, jedoch spielt diese Fertigkeit in Berufsverfahren nur eine geringe Rolle (vgl. Abele & Bühner, 2016).

Ein zweites Beispiel können Kompetenzen bei der strategischen Ausrichtung der eigenen Arbeitsgruppe bzw. der größeren Forschungs- und Lehrereinheit sowie das damit verbundene Projektmanagement genannt werden. Ist ein Bewerber/ eine Bewerberin in der Lage, langfristige Entwicklungsperspektiven zu erkennen, kann er/sie Ressourcen aktivieren, kann er/sie Kollegen und Kolleginnen gewinnen, an einer gewählten Perspektive gemeinsam zu arbeiten; kann ein/e Bewerber/in sich im (fach-) politischen Kontext bewegen und Unterstützung für bestimmte Ziele mobilisieren; sind Beispiele für entsprechende Kompetenzen.

Insgesamt sind also zwei Schwerpunkte in Berufungsverfahren zu berücksichtigen, Forschungs- und Lehrleistung sowie ergänzend soziale und Managementkompetenzen. Eine Kompensation beider Schwerpunkte sollte jedoch nicht erfolgen, da die Lehr- und Forschungsleistung immer ein bestimmendes Kriterium sein muss.

### **Verbesserte Entscheidungsgrundlagen: Diagnostik in Berufungsverfahren**

In der Forschung ist es primär die Psychologie, die Personalauswahl und Personalentwicklung optimiert. Die Psychologie etablierte im Jahre 2002 Qualitätsstandards zur berufsbezogenen Eignungsbeurteilung in Form der DIN-Norm 33430. In vielen Kursen zur DIN 33430 werden Behörden, Mittelständler, und DAX-Unternehmen darin geschult nach DIN-Norm Eignungsbeurteilungen durchzuführen. Im Grundwissen zur DIN wird darauf hingewiesen, dass anforderungsbezogen vorgegangen werden muss und die eingesetzten Instrumente die beruflichen Anforderungen möglichst objektiv, reliabel und valide abbilden sollen. Dabei ist im Vorhinein festzulegen, welche Kriterien die Personen in welcher Gewichtung erfüllen müssen.

Wir postulieren, dass die Ausführungen und Prozessschritte dieser Norm, die für Eignungsfeststellung von Führungskräften in Großunternehmen anwendbar ist, auch für die Eignungsfeststellung von Professorinnen und Professoren anwendbar sein sollte: Ein universitäres Berufungsverfahren ist ein spezifischer Fall eines Personalauswahlverfahrens. Vor der Entscheidung, welche Verfahren zur Auswahl eingesetzt werden, empfiehlt sich daher eine positionsspezifische Anforderungsanalyse. Diese kann auf der Grundlage einer Arbeits- und Anforderungsanalyse zu Aufgaben, erwarteten Ergebnissen, gewünschtem Verhalten und gewünschten Eigenschaften der auszuwählenden Person (bottom up) und /oder auf der Basis erforderlicher Kompetenzen (top down) erfolgen.

Kompetenzen im Sinne eines Top-down Ansatzes wurden im letzten Abschnitt exemplifiziert. Die folgende Liste von Verhaltensweisen kann als erste Approximation für eine Anforderungsanalyse im Sinne eines bottom-up Ansatzes verstanden werden:

- **Forschung:** Initiieren, Beantragen, Planen, Durchführen, Koordinieren, Leiten von Forschungsprojekten; Verfassen von Abschluss- und Zwischenberichten für Forschungsprojekte; Publizieren; zum Publizieren anleiten; Managen von Zeitschriften als Herausgeber oder Co-Herausgeber; Vorbereitung, Teilnahme und Leitung von Kolloquien;
- **Verwaltung:** Schreiben Gutachten für Stipendien, Abfassung von Großgeräteanträgen, Schreiben von Wiederzuweisungsanträgen für Professuren, Gutachten für diverse Gremien und (peer-review) Verfahren; Leiten von projektbezogenen Arbeitsgruppen;
- **Lehre:** Durchführung von Vorlesungen und Seminaren, Erstellung von Prüfungen, konstruktives Feedback zu Referaten geben;
- **Nachwuchsförderung:** Betreuung von Bacheloranden, Masteranden, Doktoranden und Habilitanden
- Vorbereitung, Teilnahme und Leitung von Gremien, Prüfungsausschüssen, Auswahlgremien innerhalb und außerhalb der eigenen Universität;
- **Führungskompetenz:** Führen von Mitarbeitergesprächen, Konfliktgesprächen, Schließen von Zielvereinbarungen;
- **Politisches Engagement:** Fachpolitische Aktivitäten wie Koordination von Forschungs- und Lehrbelangen über Institute hinweg; Lobbyarbeit in Bezug auf forschungs- und berufspolitische Zielsetzungen.

Auf Basis dieser (Top-down und bottom-up) Analyse können die diagnostischen Verfahren bestimmt werden, bei denen neben den klassischen Gütekriterien der einzelnen Verfahren und ihrer jeweiligen inkrementellen Validität Kosten- und Akzeptanzfragen zu beachten sind. In der Literatur werden Eigenschafts-, Simulations- und Biographie-orientierte Verfahren unterschieden.

Eigenschaftsorientierte Verfahren (z.B. Persönlichkeitstests) scheiden hier aus, da die Bewerber/innen die Tests mit hoher Wahrscheinlichkeit kennen und sich darauf vorbereiten können. Gewünschte Eigenschaften („kooperative Grundhaltung/Verträglichkeit“; „Sympathie“ – Antworten aus der Umfrage von vgl. Abele-Brehm & Bühner, 2016) müssen in anderer Weise als durch Tests diagnostiziert werden. In diesem Zusammenhang könnte z.B. das „socializing“ zu sehen sein, d.h. die Möglichkeit, Bewerber/innen nicht nur im stressreichen Bewerbungsgespräch zu erleben, sondern auch beim etwas lockereren gemeinsamen Essen (51% Zustimmung; vgl. Abele-Brehm & Bühner, 2016).

Die in Berufungsverfahren üblichen Arbeitsproben (die eingereichten Schriften, die Lehr- und wissenschaftlichen Vorträge mit ihren Fragen bzw. Diskussionsrunden mit den Studierenden und den Berufungskommissionsmitgliedern) kann man unter simulationsorientierte Verfahren subsumieren. In der Umfrage zu Kriterien bei Berufungsverfahren (Abele-Brehm & Bühner, 2016) war die Akzeptanz entsprechender Arbeitsproben (Forschungs- und Lehrvortrag; Forschungs- und Lehrkonzept; exemplarische Publikationen) hoch. Die weitergehende Verwendung von Assessment Center Übungen im engeren Sinn ist ebenfalls denkbar, derzeit aber noch wenig akzeptiert (nur 19% Zustimmung; vgl. Abele-Brehm & Bühner, 2016).

Stärkere Standardisierung der Arbeitsproben (siehe oben z.B. unsere Ausführungen zur Standardisierung des Lehrvortrags) verbessert deren Validität. In Metaanalysen zeigen standardisierte Arbeitsproben eine mittlere (.33) bis hohe (.54) kriteriumsorientierte Validität (vgl. Höft & Schuler, 2014).

Gespräche innerhalb der Berufungskommission verlaufen meist eher wenig strukturiert. Metaanalysen zeigen jedoch wiederum die eindeutigen Vorteile strukturierter (.40 - .51) gegenüber unstrukturierten Gesprächen/Interviews (.20 - .36) bei der kriteriumsorientierten Validität (klassische Studie: Schmidt & Hunter, 1996). Multimodale Interviews bieten zudem die Chance, unterschiedliche Frageformen (freie und standardisierte, biographische, situative Fragen) zu kombinieren und durch Passagen zu ergänzen, in denen die Bewerber/innen Informationen erhalten und geben. Bei der zitierten Umfrage standen zwei Drittel der Befragten der Verwendung strukturierter Interviews im Bewerbungsverfahren positiv gegenüber (Abele-Brehm & Bühner, 2016). Wir plädieren für die Entwicklung eines Leitfadens für solche strukturierten Interviews. Hierbei sollten nicht nur Fragen formuliert, sondern auch der Erwartungshorizont für Antworten spezifiziert werden. Dieser beugt Urteilerfehlern und Urteilsverzerrungen vor und führt zu einer objektiveren Bewertung. Gleichzeitig macht er deutlich, auf welchem Niveau sich die Bewerberinnen und Bewerber bei ihren Antworten bewegen und nicht nur wer sich von diesen am besten „geschlagen“ hat.

### **Qualifizierungsmöglichkeiten: Personalentwicklung**

Ist es überhaupt realistisch, eine solche vielseitig qualifizierte Person, die berühmte „eierlegende Wollmilchsau“ gewinnen zu können? Kann man davon ausgehen, dass Bewerberinnen und Bewerber auf Professuren gleichermaßen in Forschung, Lehre und überfachlichen Kompetenzen „fit“ sind, zusätzlich fachpolitisches Gespür und Engagement mitbringen und dazu möglicherweise auch noch jung, gewinnbar und bereit, den Hauptwohnsitz an den Arbeitsort zu verlegen? Die Antwort auf diese Frage lautet einerseits, dass es bei Berufungsverfahren ja darum geht, die jeweils beste Person für

diese Position zu bestimmen. Dies ist kein absolutes Maß, sondern ein relatives, bezogen auf den Kreis der Bewerbungen. Wichtiger erscheint uns jedoch eine zweite Antwort: Personaldiagnostik im Allgemeinen und im hier Besonderen dient nicht nur der Auswahl, sondern auch der Entwicklung von Bewerberinnen und Bewerbern. Wir schlagen deshalb vor, dass es für Personen, die man gern auf eine Professur berufen möchte, auch Personalentwicklung gibt. Hat man beispielsweise festgestellt, dass eine Person vom Forschungs- und Lehrprofil her besonders geeignet erscheint, sie auch kommunikativ ist, aber noch wenig Managementenerfahrung mitbringt, warum dann nicht „Qualifizierung“ bzw. „Personalentwicklung“ in Form von spezifischen Trainings? Entsprechende Personalentwicklungsmaßnahmen könnten in Zielvereinbarungen festgehalten werden – zum Nutzen aller Beteiligten. Gleiches gilt für die Lehre: sind Lehrevaluationen bzw. der Lehrvortrag nicht befriedigend, dann kann ebenfalls eine „Personalentwicklung“ für hochschuldidaktische Kompetenzen geboten sein. Entsprechende Angebote gibt es an vielen Universitäten und auch z.B. vom Hochschullehrerverband. Universitäten sollten für solche Weiterbildungsmaßnahmen Budgets bereitstellen. Dies ist in der Wirtschaft und teilweise auch in öffentlichen Verwaltungen bereits gut etablierter Standard.

### **Und wie soll das alles gehen? Wer soll das machen?**

In der Praxis scheidet ein aufwendig konzipiertes Berufungsverfahren an verfügbaren diagnostischen Instrumenten, an der Zeit und potentiell auch an der Motivation mancher Kommissionsmitglieder.

Häufig wird auch das Argument vertreten, der hier beschriebene hohe Aufwand sei nicht nötig, da die Grundquote guter Bewerberinnen und Bewerbern hoch sei. Das mag sein, jedoch wird eine auf multiplen validen Kriterien beruhende Rangfolge der zu berufenden Personen das Auswahlverfahren sicher verbessern, denn dies ist selbst bei hoher positiver Grundquote und gleichzeitig geringer Selektionsquote mit Verfahren geringerer Validität der Fall. Betrachtet man den Aufwand, den z.B. renommierte amerikanische Universitäten in Berufungsverfahren betreiben, mit dem bei uns betriebenen Aufwand, dann werden deutliche Diskrepanzen sichtbar: An diesen Universitäten werden die Bewerberinnen und Bewerber nicht nur gründlicher „getestet“, sondern sie werden auch als potentielle Kooperationspartner von morgen zuvorkommend behandelt und man widmet ihnen Zeit (vgl. Frensch, 2004). Auch wir sollten Bewerber/innen auf Stellen an unseren Universitäten entsprechend behandeln.

Es ist eine Reihe von strukturellen Maßnahmen denkbar, die eine Revision und Erweiterung von diagnostischen Maßnahmen in Berufungsverfahren eröffnen und die gleichzeitig *nicht* mit einer erhöhten Arbeitsbelastung der Kommissionsmitglieder verbunden sind:

- Es könnten zentral Leitfäden für strukturierte Interviews und für strukturierte Arbeitsproben entwickelt werden, die den Berufungskommissionen zur Hand gegeben und dann vor Beginn des Verfahrens an die Erfordernisse der jeweiligen Stelle angepasst werden.
- Die Ausarbeitung solcher strukturierter Interviews (einschließlich des Erwartungshorizonts bezüglich der Antworten) kann durch Kollegen/innen vorgenommen werden, die in Diagnostik, Personalauswahl und in Hochschulpolitik versiert sind.
- Das strukturierte Interview sollte in einer begrenzten Zeit durchführbar sein.
- Es sollte einen Leitfaden für die Auswertung und Gewichtung der Antworten geben.
- Bestimmte Kolleg/innen könnten zentral geschult werden, in Anwesenheit aller Mitglieder der Berufungskommission strukturierter Interviews durchzuführen. Dies sollten Personen sein, die sich freiwillig melden und als Berichterstatter fungieren.

- Die Auswertung der Antworten kann in Kooperation von Berufungskommission und Bericht-erstatte unter Zuhilfenahme des Leitfadens für die Auswertung erfolgen.
- Die Anzahl der Mitglieder einer Berufungskommission könnte begrenzt werden, wenn die Entscheidungsprozesse auch für Nichtmitglieder klar nachvollziehbar sind.
- Perspektivisch könnte auch an den Einsatz von Assessment Centern gedacht werden.

Zusammenfassend plädieren wir für Berufungsverfahren, bei denen die Bedeutung quantitativer Forschungsindikatoren zugunsten eines multimodalen und stärker standardisierten Verfahrens der Personalauswahl zurückgefahren wird. Nicht nur eine validere Personalauswahl, sondern auch ein ver-ringerter Publikationsdruck auf Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler sprechen dafür.

Wie in der Wirtschaft auch spricht aus unserer Sicht nur ein einziges Argument gegen einen solchen Prozess: Die zunehmende Objektivität vermindert den Spielraum für subjektive Präferenzen – auch das muss im Einzelfall bedacht werden.

#### Literatur:

Abele-Brehm, A. & Bühner M. (2016). Wer soll die Professur bekommen? Eine Untersuchung zur Be-wertung von Auswahlkriterien in Berufungsverfahren der Psychologie. *Psychologische Rundschau* \*\*\*

Frensch, P. (2004). Verschiedene Berufungskulturen, verschiedene Welten. In A. Kämmerer & J. Funke (Hrsg.), *Seelenlandschaften – Streifzüge durch die Psychologie. 98 persönliche Positionen* (S. 176-177). Göttingen: Verlag Vandenhoeck & Ruprecht.

Höft, S. & Schuler, H., (2014). Personalmarketing und Personalauswahl. In H. Schuler & K. Moser (Hrsg.). *Lehrbuch der Organisationspsychologie* (5. Aufl., S. 55–126). Bern: Huber.

Schmidt, F. & Hunter, J. (1996). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

Schönbrodt, F., Heene, M., Maier, M. & Zehetleitner, M. (2015). *Changing hiring and tenure criteria: The final jigsaw piece to incentivize credible and replicable research*. Arbeitspapier, LMU München.

Simmons, J., Nelson, L., Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359-1366.

Ulrich, R., Erdfelder, E., Deutsch, R., Strauß, B., Brüggemann, A., Hannover, B., Tuschen-Caffier, B., Kirschbaum, C., Blickle, G., Möller, J. & Rief, W. (2016). Inflation von falsch-positiven Befunden in der psychologischen Forschung: Mögliche Ursachen und Gegenmaßnahmen. *Psychologische Rundschau*, in Druck.

Wegge, J. & Rosenstiel, L. v., (2014). Führung. In H. Schuler & K. Moser (Hrsg.). *Lehrbuch der Organi-sationspsychologie* (5. Aufl., S. 315–367). Bern: Huber.

Wolbring, T. (2013). *Fallstricke der Lehrevaluation: Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität*. Frankfurt a. M.: Campus.