



DGS-Stellungnahme zum U-Multirank und CHE-Ranking¹

Stand: 17.12.2019

Die Deutsche Gesellschaft für Soziologie hat sich in den letzten Jahren mehrfach kritisch zum CHE-Ranking geäußert. Im Vergleich zum CHE-Ranking geht das von der EU initiierte U-Multirank² differenzierter vor. Nach eigenen Angaben werden beim U-Multirank 1.700 Hochschulprofile aus 96 Ländern, 5.000 Fakultäten und 11.400 Studiengängen einbezogen. Erhoben werden über 100 Indikatoren und Merkmale der beteiligten Hochschulen sowie der Fächer bzw. Fachbereiche. Diese Multidimensionalität ist zunächst positiv zu bewerten. Die Indikatoren decken die Bereiche Lehre, Forschung, Transfer, Internationalisierung und regionales Engagement ab, wobei sowohl Ranking-Indikatoren als auch deskriptive Merkmale berücksichtigt werden. Diese Felder werden sowohl durch Informationen auf der Institutionenebene als auch durch subjektive Bewertungen erfasst. Hinzu kommt die Sammlung von profildgebenden und allgemeinen Merkmalen der Institutionen, die zur Auswahl der darzustellenden Einrichtungen genutzt werden.

Die Daten zur Beschreibung dieser Merkmale werden auf unterschiedliche Weise gesammelt. Wichtige Datenquellen sind Fragebögen, die von den Hochschulen ausgefüllt werden, die Befragung von Studierenden an den beteiligten Hochschulen sowie die Auswertung der bibliometrischen Datenbank *Web of Science*.

STUDIERENDENBEFRAGUNG

Die Güte der in den Hochschulfragebogen erfassten Daten hängt grundsätzlich von der Aktualität und der Vollständigkeit der zentral verfügbaren Informationen ab. Zudem müssen surveymethodologische Aspekte in der Frage der Repräsentativität betrachtet werden. Beim U-Multirank ist unklar, ob es sich um eine zufällige Auswahl der Studierenden handelt, wie groß die Anzahl der befragten Studierenden pro Fach und Hochschule ist und ob die Unterschiede auf Signifikanz geprüft worden sind.

Die Qualität der Studierendenbefragung ist deshalb nach wie vor als problematisch anzusehen. Wenn jede teilnehmende Universität, wie beschrieben, etwa 500 Studierende pro *subject area* anschreibt und gleichzeitig bei der aktuellen Erhebung etwa 5.000 Fakultäten/Departments partizipieren, würde hieraus eine Zahl von etwa 2,5 Mio. angefragten Studierenden resultieren. Da hiervon, wie angegeben, etwa 100.000 antworten, handelt es sich um eine Rücklaufquote von lediglich 4% (also etwa 20 Fälle pro *subject area*) – trotz der möglicherweise erfolgten selektiven Auswahl und Ansprache der

¹ Basierend auf den Ergebnissen einer im September 2019 eingesetzten Task Force, bestehend aus Jörg Blasius (Bonn), Nicole Burzan (Dortmund), Monika Jungbauer-Gans (Hannover), Stephan Lessenich (München) und Uwe Vormbusch (Hagen).

² Zur Zielsetzung und Methode von U-Multirank siehe: www.che-ranking.de/cms/?getObject=1176&getLang=de; Zugriff 28.11.2019

Studierenden durch die jeweiligen Departments. Bei einer solch niedrigen Rücklaufquote bzw. geringen Fallzahl pro *subject area* kann keinesfalls davon ausgegangen werden, dass die Antwortenden ein auch nur ungefähres Abbild der Grundgesamtheit darstellen. Somit ist festzuhalten, dass anhand der eingeholten studentischen Perspektiven keine zuverlässigen Aussagen getroffen werden können. Informationen darüber, ob die Universitäten jeweils die Verteilung bestimmter Merkmale der Studierenden in den *subject areas* an U-Multirank liefern, die mit der Struktur der befragten Studierenden verglichen werden, sind in der Selbstdarstellung des U-Multirank nicht zu finden.

Bei den studentischen Beurteilungen besteht außerdem das gleiche Problem wie auch bei Lehrveranstaltungsevaluationen generell: Zufriedenheit der Studierenden und Güte des Studienangebots werden gleichgesetzt, Störvariablen bzw. Grundlagen der Erwartungen und Beurteilungen und dergleichen werden nicht berücksichtigt. Unklar bleibt zudem größtenteils, wie die Antworten in ein Ranking übersetzt, wie Einzelaspekte gewichtet und wo Grenzen in den Abstufungen gezogen werden.

Nach wie vor besteht bei der Studierendenbefragung außerdem das Grundproblem, dass lokale Bewertungen miteinander verglichen und zumindest teilweise auch in eine Rangordnung gebracht werden. Die Befragten haben in der Regel allerdings nur von ihrem eigenen lokalen Kontext Kenntnis, der daher den alleinigen Bewertungshorizont bildet. Der durch das U-Multirank angestrebte Vergleich von Universitäten auf dieser Basis ist nicht valide.

BIBLIOMETRISCHE DATEN

Bezüglich der bibliometrischen Daten ist festzuhalten, dass die im *Web of Science* erfassten Journals spezifische Verzerrungen aufweisen. Vor einer Auswertung sind umfangreiche Bereinigungen durchzuführen, wobei auf den Vorarbeiten des Centre for Science and Technology Studies (CWTS) für das Leiden Ranking aufgebaut wird. In dieser Datenbasis sind nicht alle Fächer gleichermaßen erfasst, auch die englischsprachigen referierten Journals überwiegen gegenüber Journals in anderen Landessprachen. Da sich Publikationskulturen unterscheiden, können Fächervergleiche nicht erfolgen. Bekanntlich unterscheiden sich zum Teil sogar die Publikationskulturen innerhalb von Fächern (wie beispielsweise auch in der Soziologie), so dass unterschiedliche Fächerspezialisierungen an verschiedenen Hochschulen zu unterschiedlichen Ergebnissen bei bibliometrischen Analysen führen.

Bei der Bemessung des Publikationsoutputs stellt sich außerdem die Frage, welche An-Institute und lose verbundene Einheiten mitgerechnet werden. In Bezug auf die Auswertung von Patentanmeldungen müssten auch Fächerprofile kontrolliert werden.

ERGEBNISPRÄSENTATION

Bei der Darstellung der Ergebnisse können sehr flexibel Merkmale für den Vergleich ausgewählt werden. Es besteht die Möglichkeit, die gesamte Institution oder einzelne Fächer zu vergleichen. Man kann sich fast beliebig Profile auswählen, was gut ist, wenn man nach bestimmten Merkmalen sucht. Schränkt man die Suche auf einzelne Länder und Fächer ein, sieht man, dass die Institutionen recht (selbst)selektiv sind, d.h. es beteiligen sich oft nicht viele Hochschulen, so dass man die Datenbasis nicht wie vorgesehen als Quelle für systematische Vergleiche nutzen kann.

Auf der U-Multirank-Seite sind die Länderbeschreibungen informativ – als Orientierung für ausländische Studierende. Dies scheint die hauptsächliche Zielsetzung von U-Multirank zu sein. Für mobilitätsinteressierte ausländische Studierende könnte eine solche Datenbasis eine Quelle sein, um sich ein Bild vom Zielland zu machen und anschließend eine Hochschule, z.B. mit hoher Forschungsleistung, auszusuchen. Für die typischen inländischen Studierenden, die mit einem Bachelor anfangen wollen, werden viele Kriterien nicht zentral sein. Hier würden eher differenzierte inhaltliche Profile interessieren, die allerdings nicht vorhanden sind.

RANGLISTEN

In der partiellen Personalisierbarkeit von Vergleichsindikatoren und Darstellungsmöglichkeiten drückt sich im Vergleich zum CHE-Ranking ein facettenreicheres und auf Variabilität der Vergleichsmaßstäbe angelegtes Verständnis von Evaluation aus. Der Umstand, dass man in den Ergebnistabellen die Universitäten nach wie vor nach „top scores“ sortieren kann (wobei offensichtlich mangels Datengrundlage unbesetzte Felder nicht zu einem Punktabzug führen), unterläuft allerdings die eingangs angesprochene größere Differenzierung. Das Bekenntnis zur Multidimensionalität hält U-Multirank zudem nicht davon ab, auf der Homepage die „top 25 performing universities“ weltweit nach 10 Kriterien auszuweisen. Dass es weiterhin Präsentationsformate wie die „Olympiatabelle“ gibt, in der Universitäten ausschließlich auf der Basis vergebener „A-scores“ konkurrieren, verweist sowohl auf die Persistenz der Vorstellung von Höchstleistungsstandorten, für die eigene Regeln gelten, als auch auf eine der Komplexität der Situation nicht gerecht werdende Simplifizierungsrhetorik.

WISSENSCHAFTSPOLITISCHE RELEVANZ

In der wissenschaftlichen Bewertung lässt sich grundsätzlich darüber diskutieren, inwiefern Rankings dem Risiko einer einseitigen Quantifizierung von Bewertungskriterien Vorschub leisten, die die Güte von Forschung – auf dieser Ebene inhaltsabstinent – nur noch sehr indirekt einschätzen können. Es lässt sich allerdings nicht bestreiten, dass z.B. die quantitative Publikationsaktivität bzw. Zitationen eine hohe Bedeutung für die Beurteilung wissenschaftlicher Qualität haben. Sie müssen jedoch zu anderen Kriterien in ein ausgewogenes Verhältnis gesetzt werden.

EMPFEHLUNG

Die von der DGS in den Jahren 2012 und 2015 geäußerten methodischen und fachpolitischen Bedenken zum CHE-Ranking bleiben somit auch im Fall des U-Multiranks, trotz festzustellender Verbesserungen, weitgehend bestehen. Dies gilt z.B. für die Qualität der Studierendenbefragung und der hierdurch erhobenen Daten, für die Scheinneutralität verschiedener gewählter Indikatoren (z.B. Wissenstransfer), oder für die grafische Präsentation und öffentlichkeitswirksame Darstellung der Beurteilungen, die auch weiterhin einer simplifizierenden Interpretation Vorschub leistet. Inhaltliche Informationen zu den Studiengängen stehen hier im Gegensatz zu dem von den Fachgesellschaften entwickelten Studieninformationssystem *studium.org*/ (www.studium.org) nicht im Fokus des Interesses.

Die wohl begründete Empfehlung der DGS an die Institute lautet deshalb, auf eine Beteiligung am U-Multirank (wie auch weiterhin am CHE-Ranking) zu verzichten und Interessierte stattdessen in geeigneter Weise auf studium.org/soziologie zu verweisen.