

Internationale Hochschulrankings Ein methodenkritischer Vergleich

Anna Kroth · Hans-Dieter Daniel

Zusammenfassung: Trotz ihrer umstrittenen Validität, Zuverlässigkeit und Wirkung, werden internationale Universitätsrankings immer häufiger von Studierenden und Universitäten als Informationsquellen genutzt. Dieser Artikel vergleicht drei internationale Hochschulrankings und analysiert, basierend auf der wissenschaftlichen Literatur, inwiefern sie relevante Aspekte von Forschung oder Lehre erheben und diese adäquat und verlässlich messen. Untersucht wurden das *Academic Ranking of World Universities (ARWU Ranking)*, das *World University Ranking* und das vom deutschen Centrum für Hochschulentwicklung (CHE) entwickelte *Excellence Ranking* für Master- und Doktorandenprogramme. Die Analyse basiert auf den Publikationen der Herausgeber der Rankings und auf wissenschaftlichen Artikeln über diese internationalen Rankings. Basierend auf der Analyse werden Schlussfolgerungen für die Gestaltung internationaler Rankings abgeleitet und weiterführende Forschungsfragen formuliert.

Schlüsselwörter: Internationales · Hochschulranking · Universität · Reliabilität · Validität

International University Rankings – A Critical Review of the Methodology

Abstract: Despite disagreement about their validity, reliability and effects, international rankings of universities have become widely used by students and higher education institutions. This article compares three international rankings of universities and analyzes, based on the academic literature, the extent to which they assess relevant aspects of teaching and research and whether these aspects are measured adequately. We study the *Academic Ranking of World Universities* published by Shanghai Jiao Tong University, the *World University Ranking* published by the British magazine *Times Higher Education Supplement* and the *Excellence Ranking* published by the German Centre for Higher Education Development (CHE). The analysis is based on the academic literature about these international rankings and the publications by the rankings' publishers. We draw conclusions about the design of rankings and develop further research questions.

Keywords: Higher education · International ranking · Reliability · University · Validity

1. Internationale Hochschulrankings

Innerhalb weniger Jahre sind internationale Hochschulrankings zu einer häufig verwendeten Informationsquelle avanciert. In Studien aus England nannten beispielsweise 29 % der Studienanfänger nationale und internationale Universitätsrankings als wichtige Informationsquelle (UNITE 2007) und die Mehrheit der Universitäten gab an, sie für strategische Entscheidungen zu nutzen (Hazelkorn 2007). Doch trotz ihrer weiten Verbreitung sind die Methodik und die Wirkungen internationaler Hochschulrankings sehr umstritten. Befürworter argumentieren, dass Rankings, auch wenn sie nur einen Ausschnitt der Leistungen von Universitäten widerspiegeln, Studierenden bei der Wahl einer geeigneten Universität helfen, Universitäten zusätzliche Anreize zur Leistungssteigerung geben und ihnen wertvolle Informationen für Verbesserungsbemühungen zur Verfügung stellen.

Diesen Argumenten wird häufig entgegengehalten, dass die bestehenden internationalen Universitätsrankings der Forschung und Lehre eher schaden als nutzen, weil sie wichtige Aspekte der Qualität von Universitäten nicht erheben, zu simple Indikatoren verwenden und zahlreiche Messfehler aufweisen. Es wird befürchtet, dass solche Rankings zu einer Anpassung der Hochschulen an die Kriterien der Rankings führen. Leistungen, die nicht adäquat in Rankings gemessen werden, wie zum Beispiel die Lehrqualität, die geisteswissenschaftliche Forschung oder gesellschaftliches Engagement könnten daher vernachlässigt werden (Dill/Soo 2005; Marginson/Van der Wende 2007). Auch könnte eine stärkere Vereinheitlichung der Universitäten zum Typ der großen Forschungsuniversität eine unintended Folge von Universitätsrankings sein. Eine weitere Befürchtung ist, dass Universitätsrankings zu einer stärkeren vertikalen Stratifizierung zwischen Universitäten führen. Oftmals betonen Rankings geringe oder vermeintliche qualitative Unterschiede, was jedoch für Universitäten auf den niedrigeren Rangplätzen einen schweren Nachteil im Wettbewerb um Ressourcen bedeutet und ihre Absolventen benachteiligt (Espeland/Sauder 2007).

In diesem Artikel konzentrieren wir uns auf die Methodik internationaler Hochschulrankings und analysieren ihre Konstruktvalidität, Messvalidität und Reliabilität. Wir erforschen, inwiefern diese Rankings die betreffenden Leistungen in Forschung oder Lehre adäquat und verlässlich abbilden. Untersucht werden das von der Shanghai Jiao Tong University publizierte *Academic Ranking of World Universities (ARWU Ranking)*, das von der britischen Zeitung Times Higher Education Supplement publizierte *World University Ranking (THES Ranking)* und das vom deutschen Centrum für Hochschulentwicklung entwickelte *CHE Excellence Ranking*. Die Analyse basiert auf den Publikationen und Webseiten der Herausgeber dieser Rankings und auf wissenschaftlichen Artikeln über diese internationalen Rankings. Aus dieser Analyse leiten wir Schlussfolgerungen für die Gestaltung internationaler Rankings ab.

Gelegentlich wird die Forschung über die Rankingmethodik mit dem Argument kritisiert, dass Rankings auch mit einer verbesserten Methodik die oben genannten negativen Auswirkungen zur Folge hätten. Wir stimmen zu, dass verzerrende Anreize und eine wachsende Ungleichheit zwischen Universitäten auch bei methodisch verbesserten Rankings entstehen können, wenn auch in einem geringeren Ausmaß. Diese nicht intendierten Auswirkungen sollten dringend untersucht werden. Trotzdem erachten wir, genauso wie eine Anzahl weiterer Hochschulforscher, es als wichtig, die Rankingmethodik zu untersuchen. In

den letzten Jahren ist zunehmend eine globale „audit culture“ (Strathern 2000) entstanden, in der *accountability* und Qualitätsvergleiche so hoch geschätzt werden, dass es unmöglich erscheint, Hochschulrankings zu verhindern. Die Forschung über Rankingmethodik und über Möglichkeiten zur Erhöhung ihrer Validität und Reliabilität ist daher wichtig, um die negativen Auswirkungen zu begrenzen und mögliche positive Effekte zu bewirken (Altbach 2006; Van Raan 2005; Dill/Soo 2005; Marginson/Van der Wende 2007).

2. Internationale Universitätsrankings im Überblick

In den letzten Jahren sind eine Vielzahl internationaler Universitätsrankings publiziert worden. Wir haben die oben genannten Rankings für die Analyse ausgewählt, weil in ihnen mehrere Fachgebiete oder gesamte Universitäten beurteilt und sowohl Forschung als auch Lehre betrachtet werden. Ein Vergleich der Rankings zeigt, dass in ihnen trotz ähnlicher Zielsetzungen unterschiedliche Qualitätskonzepte, Indikatoren und Gewichtungen verwendet werden. Die Charakteristika der Rankings haben wir tabellarisch zusammengefasst (vgl. Tabelle 1).

2.1 Das ARWU Ranking der Shanghai Jiao Tong University

Das *ARWU Ranking* wurde von Hochschulforschern der Shanghai Jiao Tong University entwickelt, um die Qualität chinesischer Universitäten im internationalen Vergleich einzuschätzen und die Ergebnisse für Verbesserungsprozesse zu nutzen (Liu/Cheng 2005a, b). Es wird seit dem Jahr 2003 veröffentlicht. Die Rangplätze werden nach folgenden Kriterien bestimmt (SJTU 2007): Die Qualität der Ausbildung wird anhand der Anzahl der Alumni gemessen, die einen Nobelpreis oder eine *Fields Medal* erhalten haben. Dieser Wert geht mit 10 % in die Bewertung ein. Die Qualität der Lehrenden wird mittels der Anzahl der Lehrenden, die einen Nobelpreis oder eine *Fields Medal* erhalten haben, und der Anzahl der *Highly Cited Researchers* in 21 Fächergruppen ermittelt, die in der Literaturdatenbank *Web of Knowledge*¹ von Thomson Reuters verzeichnet sind, und geht mit jeweils 20 % in die Bewertung ein. Der Forschungs-Output wird mittels der Anzahl der Artikel, die in den Zeitschriften *Nature* und *Science* publiziert werden und im *Science Citation Index – expanded* bzw. im *Social Sciences Citation Index* gelistet sind, ermittelt und geht mit jeweils 20 % in die Bewertung ein. Um kleinere Universitäten nicht zu benachteiligen, wird das Ergebnis mit der Anzahl der vollzeitäquivalenten akademischen Mitarbeiter gewichtet. Dieser Wert geht mit 10 % in das Ergebnis ein. Das *ARWU Ranking* bewertet gesamte Universitäten. Im Jahr 2008 wurde darüber hinaus ein nach ähnlicher Methodik gestaltetes fächergruppenspezifisches Universitätsranking publiziert (SJTU 2008).

2.2 Das THES Ranking

Das *THES Ranking* wurde von der britischen Zeitung *Times Higher Education Supplement* entwickelt und erstmals 2004 publiziert. Laut den Herausgebern verfolgt das Ranking primär das Ziel, international mobile Studierende bei der Wahl einer Universität zu unter-

Tabelle 1: Übersicht über internationale Hochschulrankings (Quellen: SJTU 2007; THES 2007; CHE 2007)

| | <i>ARWU Ranking, Shanghai Jiao Tong University</i> | <i>THE Ranking, Times Higher Education Supplement</i> | <i>CHE Excellence Ranking, Centrum für Hochschulentwicklung</i> |
|-------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Analyseebene | • Gesamte Hochschule | • Gesamte Hochschule | • Fachspezifisch |
| Präsentation der Ergebnisse | <ul style="list-style-type: none"> • Rangliste • Indikatoren werden aggregiert | <ul style="list-style-type: none"> • Rangliste • Indikatoren werden aggregiert | <ul style="list-style-type: none"> • Bildung einer Exzellenz-Gruppe; innerhalb derer werden Ranggruppen für überdurchschnittliche Ergebnisse gebildet (Gold-, Silber- oder Bronze-Medaillen) • Jeder Indikatorwert wird aufgeführt |
| Outputs und Outcomes der Forschung | <ul style="list-style-type: none"> • Anzahl der in <i>Nature</i> und <i>Science</i> publizierten Artikel zwischen 2002-2006. Gewicht: 20 % • Anzahl der im <i>Science Citation Index - expanded</i> und im <i>Social Sciences Citation Index</i> indexierten Artikel im Jahr 2006 20 % • Anzahl der Forscher, die zwischen 1911 und 2006 einen Nobelpreis oder die <i>Fields Medal</i> gewonnen haben 20 % • Anzahl der <i>Highly Cited Researchers</i> zwischen 1981-1999 in 21 Fachgebieten der Lebenswissenschaften, Medizin, Physik, Ingenieur- und Sozialwissenschaften 20 % • Gewichtung für Hochschulgröße, indem das Ergebnis mit der Anzahl der Forscher gewichtet wird 10 % | <ul style="list-style-type: none"> • Zahl der Zitationen pro Forscher von 2002 bis 2007, basierend auf Angaben aus der SCOPUS-Literaturdatenbank 20 % • Durchschnitt der pro Publikation erreichten Zitationen gewichtet mit der durchschnittlichen Zitierhäufigkeit in dem Fachgebiet • Anzahl der <i>Highly Cited Researchers</i> 1984-2003, Nobelpreisträger und Gewinner der <i>Fields Medal</i> • Anzahl der geförderten Projekte im Marie-Curie-Programm • Für die Fachbereiche der Exzellenz-Gruppe werden die Ergebnisse von Studierendenbefragungen und die Forschungsschwerpunkte publiziert | <ul style="list-style-type: none"> • Anzahl der Publikationen der Vertreter des Faches in der Datenbank <i>Web of Science</i> 1997-2004 • Durchschnitt der pro Publikation erreichten Zitationen gewichtet mit der durchschnittlichen Zitierhäufigkeit in dem Fachgebiet • Anzahl der <i>Highly Cited Researchers</i> 1984-2003, Nobelpreisträger und Gewinner der <i>Fields Medal</i> • Anzahl der geförderten Projekte im Marie-Curie-Programm • Für die Fachbereiche der Exzellenz-Gruppe werden die Ergebnisse von Studierendenbefragungen und die Forschungsschwerpunkte publiziert |

Tabelle 1 Fortsetzung

| | <i>ARWU Ranking, Shanghai Jiao Tong University</i> | <i>THE Ranking, Times Higher Education Supplement</i> | <i>CHE Excellence Ranking, Centrum für Hochschulentwicklung</i> |
|--------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Output und Outcomes der Lehre | <ul style="list-style-type: none"> Anzahl der Alumni, die zwischen 1911 und 2006 einen Nobelpreis oder die <i>Fields Medal</i> gewonnen haben 10 % | <ul style="list-style-type: none"> Zahlenverhältnis zwischen Lehrenden und Studierenden 20 % Anteil international mobiler Studierender 5 % Anteil ausländischer Forscher 5 % | <ul style="list-style-type: none"> Zahlenverhältnis zwischen Lehrenden und Studierenden 20 % Anteil international mobiler Studierender 5 % Anteil ausländischer Forscher 5 % |
| Reputation der Hochschule | | <ul style="list-style-type: none"> 3703 Forscher nennen die 30 führenden Hochschulen ihres Gebiets, gewichtet nach Regionen und Fachgebieten 40 % 736 internationale Arbeitgeber wurden gefragt, an welchen Hochschulen die besten Absolventen studiert haben 10 % | |

stützen (THES 2007). Das *THES Ranking* bewertet Universitäten zu einem großen Teil nach ihrer internationalen Reputation. Vierzig Prozent der Bewertung basieren auf einer Umfrage, in der Professoren aus allen Kontinenten gebeten werden, die besten Universitäten ihres Gebiets zu nennen. Weitere 10% basieren auf einer Meinungsumfrage von internationalen Arbeitgebern. Daneben wird das Zahlenverhältnis zwischen Lehrenden und Studierenden als Indikator für die Qualität der Lehre verwendet. Dies macht 20 % des Rangingergebnisses aus. Weitere 20% entfallen auf die Anzahl der Zitationen in der Literaturdatenbank *SCOPUS* pro akademischen Mitarbeiter. Dies wird als Indikator für die Qualität der Forschung gesehen. Des Weiteren betrachtet *THES* den Grad der Internationalisierung. Dieser wird anhand des Anteils der international mobilen Studierenden und Lehrenden gemessen und geht mit jeweils 5 % in die Bewertung ein (THES 2007).

2.3 Das Excellence Ranking des Centrums für Hochschulentwicklung

Das *Excellence Ranking* wurde 2007 vom Centrum für Hochschulentwicklung (CHE) entwickelt. Es soll Studierende bei der Wahl geeigneter Master- oder Doktorandenprogramme unterstützen, Universitäten als Marketingwerkzeug dienen und ein Benchmarking von Universitäten ermöglichen. Außerdem soll das Ranking Förderinstitutionen und die Öffentlichkeit über die Leistungen der Universitäten informieren (CHE 2007). Im Unterschied zu den anderen beiden Rankings werden im *Excellence Ranking* keine Ranglisten, sondern Ranggruppen publiziert und die Leistungen der Universitäten anhand von vier Indikatoren dargestellt, die nicht zu einem Index zusammengefasst werden. Für herausragende Leistungen werden für jeden Indikator Gold-, Silber- oder Bronzemedailen verteilt. Es wird ein zweistufiges Verfahren verwendet, um die Medallengewinner zu bestimmen. Zunächst werden europäische Universitäten nach vier Indikatoren beurteilt. Im zweiten Schritt werden die Universitäten, die zumindest in drei der vier Indikatoren eine Gold- oder eine Silbermedaille erhalten haben, genauer untersucht. Zu diesem Zweck werden die Hochschulen, die Masterstudierenden und Postgraduierten mittels standardisierter Fragebögen befragt (Berghoff/Brandenburg/Müller-Böling 2008). Das Ranking präsentiert die Ergebnisse fachspezifisch und hat in der ersten Ausgabe Studiengänge in Biologie, Chemie, Physik und Mathematik untersucht. In der nahen Zukunft plant das CHE, weitere Fächer einzubeziehen (Berghoff/Brandenburg/Müller-Böling 2008).

Zur Bewertung wurden im ersten Schritt anhand von vier Indikatoren (Anzahl der Publikationen der Fachvertreter an der jeweiligen Institution in der Literaturdatenbank *Web of Science*, Durchschnitt der pro Publikation erreichten Zitierungen – gewichtet mit der durchschnittlichen Zitierhäufigkeit des Fachgebiets –, Anzahl der von Thomson Reuters angegebenen *Highly Cited Researchers* und der Anzahl der Projekte im Marie-Curie-Förder-Programm) Hochschulen in Europa identifiziert, die zumindest in einem dieser Kriterien besser als der Durchschnitt waren und dafür eine Bronzemedaille oder bei besseren Leistungen eine höherwertige Medaille erhielten. Hochschulen, die mindestens eine Silbermedaille in drei der vier Indikatoren erringen konnten, wurden der Exzellenz-Gruppe zugeordnet. Für diese Exzellenz-Gruppe wurden unter anderem Daten zur Zufriedenheit der Studierenden und zu den Forschungsschwerpunkten der Universitäten erhoben.

Die oben dargestellten internationalen Universitätsrankings gehören zu den Bekanntesten. Jedoch können viele weitere Aspekte, wie beispielsweise die Gleichstellung von Frauen

oder die extracurricularen Angebote, in einem Universitätsranking verglichen werden. Für den Bildungssektor der USA gibt es über 4600 verschiedene Rankings (Hattendorf Westney 2006). Beispiele für weitere internationale Universitätsrankings sind das *Newsweek Top 100 Global Universities Ranking* (Newsweek 2006), das die Ergebnisse des *ARWU Rankings* und des *THES Rankings* kombiniert und Reputationswerte ausschließt sowie die ausschließlich auf bibliometrischen Daten basierenden *Leiden Rankings*, die von Wissenschaftlern des niederländischen Centers for Science and Technology Studies (CWTS 2007) entwickelt wurden. Ihre Rankings beurteilen Universitäten unter anderem nach der durchschnittlichen Zitierhäufigkeit der Veröffentlichungen normiert mit der durchschnittlichen Zitierhäufigkeit des Fachgebiets (CWTS 2007). Daneben existiert eine Vielzahl fachspezifischer Rankings. Für MBA Studiengänge hat die Financial Times (2008) ein Ranking erstellt und für volkswirtschaftliche Fachbereiche existiert ein Ranking der Europäischen Kommission (2004). Daneben gibt es Universitätsrankings, die auf seltener genutzten Informationen basieren, wie zum Beispiel das von der Information Science and Scientometrics Research Unit (ISSRU) entwickelte Ranking, dessen Bewertung auf der Anzahl der Universitätsmitglieder in den Editorial und Advisory Boards von jeweils 20 ausgewählten internationalen Zeitschriften in zwölf Fachgebieten basiert (Braun et al. 2007) oder das *Webometrics Ranking of World Universities*, das den Umfang und die Vernetzung der Internetpublikationen von Universitäten berücksichtigt (Aguillo/Ortega/Fernández 2008; Cybermetrics Lab 2007).

3. Methode

Um zu analysieren, inwiefern die drei internationalen Universitätsrankings die Leistungen in Forschung und Lehre valide und reliabel messen und die in der Einleitung genannten verzerrenden Anreize ausüben, analysieren wir zuerst die Konstruktvalidität der internationalen Rankings. Dieses Gütekriterium betrifft die Frage, inwiefern ein Messinstrument das zu messende Konstrukt abbildet (Shadish/Cook/Campbell 2002). Shadish, Cook und Campell unterscheiden zwei Aspekte von Konstruktvalidität: Ein Aspekt ist, inwiefern die Indikatoren inhaltlich dem zu messenden Konstrukt entsprechen. Für unsere Analyse bedeutet dies, ob die Indikatoren wichtige Aspekte von Forschung und Lehre erfassen. Der zweite Aspekt bezieht sich auf die Frage, inwieweit die verwendeten Indikatoren adäquat gemessen werden. Der Einfachheit wegen nennen wir den ersten Aspekt Konstruktvalidität und den zweiten Messvalidität. Als weiteres Kriterium betrachten wir die Reliabilität. Sie bezieht sich auf die Frage, ob ein Messverfahren bei wiederholten Messungen die gleichen Ergebnisse produziert, ob die Ergebnisse zeitlich stabil sind und ob die Ergebnisse mit anderen Qualitätsmessungen übereinstimmen.

4. Konstrukt- und Messvalidität der internationalen Universitätsrankings

4.1 Das ARWU Ranking der Shanghai Jiao Tong University

Messprobleme der „Web of Knowledge“-Literaturdatenbanken. Ein Problem des *ARWU Rankings*, und aller Universitätsrankings die auf der *Web of Knowledge*- oder der

SCOPUS-Literaturdatenbank basieren, ist, dass diese Datenbanken nicht immer zuverlässige Informationen enthalten. Van Raan (2005) macht darauf aufmerksam, dass 7 % der zitierten Publikationen im *Web of Science* nicht identifiziert werden können und dass die Zuordnung von Publikationen zu Universitäten ebenfalls nicht immer gelingt. Insbesondere bei Universitäten, die unterschiedliche Namen verwenden, oder verschiedene Institute haben, die unter eigenem Namen publizieren, ist die Zuordnung von Publikationen schwierig. Trotz dieser Probleme gibt das *Web of Knowledge* wertvolle Informationen über Publikationsleistungen. Es ist jedoch wichtig, zu berücksichtigen, dass die Messvalidität dieser Indikatoren und der Rankings, die darauf basieren, eingeschränkt ist. Um die Verlässlichkeit des *Web of Knowledge* zu erhöhen, wäre es sinnvoll, Datenbankeinträge sorgfältiger durchzuführen und insbesondere Universitätsnamen einheitlich zu verwenden.

Mangelnde Berücksichtigung der Sozial- und Geisteswissenschaften. Eine weitere Schwäche des *ARWU Rankings* ist, dass es die Publikationsleistungen in den Sozial- und Geisteswissenschaften nicht adäquat widerspiegelt, weil diese weniger häufig in wissenschaftlichen Zeitschriften publizieren (Van Raan 2005; Marginson/Van der Wende 2007). Universitäten mit einem Schwerpunkt in den Sozial- und Geisteswissenschaften sind in diesem Ranking daher benachteiligt. Von diesem Problem sind ebenfalls alle Universitätsrankings betroffen, die auf der *Web of Knowledge*-Literaturdatenbank basieren. Eine Möglichkeit, dieses Problem zu minimieren, besteht darin, neben Zeitschriften auch andere Publikationsarten zu berücksichtigen. Zudem ist es sinnvoll, die Zitierhäufigkeit mit der durchschnittlichen Zitierhäufigkeit des Fachgebiets zu gewichten, so wie es im *Leiden Ranking* und im *Excellence Ranking* praktiziert wird. Weiterhin bevorteilt das *ARWU Ranking* Universitäten mit einem natur- und lebenswissenschaftlichen Fokus, weil es die Anzahl der Artikel in den Zeitschriften *Nature* und *Science* als Leistungsindikator verwendet (Marginson/Van der Wende 2007).

Bevorteilung großer Universitäten. Ein weiterer Schwachpunkt ist, dass das *ARWU* die Anzahl der Publikationen als Indikator für die Forschungsqualität verwendet. So wird v. a. das Publikationsvolumen und nicht notwendigerweise die Qualität bewertet. Kleinere Universitäten sind beim Vergleich absoluter Indikatorwerte im Nachteil. Die im *ARWU* verwendete Gewichtung des Ergebniswerts mit der Anzahl der vollzeitäquivalenten akademischen Mitarbeiter wird als ein zu geringer Ausgleich betrachtet, da alle anderen Indikatoren, große Forschungsuniversitäten bevorteilen (Marginson/Van der Wende 2007). Die Gewichtung der Publikations- oder Zitationszahlen nach der Größe einer Universität ist eine der Herausforderungen für internationale Universitätsrankings. Eine Möglichkeit, dieses Problem abzuschwächen, wäre die Anzahl der Publikationen oder Zitationen mit der Anzahl der Forscher zu gewichten. Das *THES Ranking* berechnet seinen Indikator für Forschungsqualität bspw. indem es die Anzahl der Zitationen mit der Anzahl der Forscher gewichtet. Jedoch ist es fast unmöglich, die Anzahl der Forscher einer Universität international vergleichbar zu erheben. Dies liegt einerseits an den internationalen Unterschieden in den Beschäftigungsformen an Universitäten. Eine Frage ist beispielsweise, ob die Lehrbeauftragten und Privatdozenten ohne Stelle an der Universität auch zu den Mitarbeitern gezählt werden sollten. Daneben ist es schwierig, die Anzahl der Forscher

festzustellen, weil Universitäten in einigen Ländern „geschönte“ Personalstatistiken präsentieren, die beispielsweise nur die forschungsaktiven Personen umfassen. In den USA ist dieses Verhalten häufig beobachtet worden (Ehrenberg 2002). Weitere Forschung über die Frage, wie bei der Beurteilung von Forschungsleistungen die Größe einer Universität adäquat gewichtet werden kann, stellt ein wichtiges Desiderat dar.

Ungeeignete Indikatoren für die Leistungen in der Lehre. Im *ARWU Ranking* wird die Anzahl der Forscher, die in der Literaturdatenbank *Web of Knowledge* als *Highly Cited Researchers* (unterteilt in 21 Fächergruppen) ermittelt werden, als Indikator für die Qualität der Lehrenden dargestellt. Dieser Indikator ist, mit den oben genannten Einschränkungen, geeignet, um die Forschungsqualität einzuschätzen. Hinsichtlich der Qualität der Lehrenden ist die Aussagekraft jedoch eher gering und nicht für eine solide Bewertung geeignet. Viel publizierende Forscher müssen nicht unbedingt gute Lehrende sein (Marsh/Hattie 2002). Bei dem Indikator *Highly Cited Researchers* ergibt sich zudem das Problem, dass die Anzahl der viel zitierten Forscher sich auf den Zeitraum zwischen 1981 bis 1999 bezieht (Thomson ISI 2008). Aktuelle Publikations- und Zitationserfolge der letzten Jahre werden somit nicht berücksichtigt.

Die Anzahl der Nobelpreisträger ist nur für wenige Fächer aussagekräftig. Im *ARWU Ranking* wird die Anzahl der Forscher, die einen Nobelpreis oder eine *Fields Medal* erhalten haben, als Indikator für die Qualität der Forschung verwendet. Dieser Indikator hat nur Aussagekraft für wenige Fächer und Universitäten. Für die Qualität der Lehre kann die Anzahl der ehemaligen Studierenden, die viele Jahre später einen Nobelpreis erhalten, kein Leistungsindikator sein. Messprobleme ergeben sich bei der Operationalisierung des Konstrukts (Braun/Szabadi-Peresztegi/Kovács-Németh 2003) und bei der Zuordnung von Preisträgern zu Universitäten (Van Raan 2005).

Insgesamt hat das *ARWU Ranking* eine vergleichsweise hohe Konstruktvalidität für die Bewertung von Forschungsleistungen in den Natur- und Lebenswissenschaften (Marginson/Van der Wende 2007). Die Sozial- und Geisteswissenschaften werden weniger valide erhoben. Entgegen eigener Angaben, verfügt das *ARWU Ranking* über keine adäquaten Indikatoren für die Qualität der Lehre. Die Tatsache, dass das *ARWU Ranking* auf mehr oder minder objektiven Daten basiert, erhöht die Messvalidität, auch wenn diese aufgrund der Zuordnungsprobleme im *Web of Knowledge* etwas eingeschränkt ist.

4.2 Das THES Ranking des Times Higher Education Supplement

Reputationswerte sind ungeeignete Indikatoren. Das *THES Ranking* unterscheidet sich stark vom oben dargestellten Ranking, weil die Ergebnisse nicht auf objektiven Daten, sondern zu 50 % auf den Meinungen von Professoren und Arbeitgebern basieren. Es wird häufig kritisiert, dass das *THES Ranking* solche Reputationswerte als Indikator für die Qualität von Forschung und Lehre ausgibt (Taylor/Braddock 2007; Marginson/Van der Wende 2007). Reputationswerte sind keine adäquaten Indikatoren für die Qualität von Forschung und Lehre, weil der Ruf einer Universität noch lange nach der tatsächlichen Leistung anhält und traditionsreiche Universitäten bevorteilt sind. Gu-

rino et al. (2005) argumentieren daher, dass Reputationsrankings die bestehende Reputation recyceln und keine neuen Informationen bereitstellen. Reputationsmessungen bevorzugen auch die Universitäten in der jeweiligen Region der Befragten (Taylor/Braddock 2007). Marginson (2007) erklärt das außergewöhnlich gute Abschneiden der australischen Universitäten damit, dass sie den Befragten in der asiatischen Region besser bekannt sind als möglicherweise bessere Universitäten in anderen Regionen. Insbesondere in internationalen Rankings sind Reputationswerte ungeeignet, weil es wohl keiner Person möglich ist, die Qualität aller Universitäten in der Welt zuverlässig zu vergleichen. Aus gleichen Gründen liefert auch die Befragung von Arbeitgebern keine brauchbaren Aussagen über die Lehrqualitäten von Universitäten. Neben der Konstruktvalidität ist auch die Messvalidität solcher Umfragen sehr zweifelhaft. Die Bitte an die Professoren lautete: „Comment on universities in your area“. Diese Frage ist zweideutig, da unklar ist, ob die besten Universitäten im eigenen Fachgebiet (area = Studienfach) oder in der eigenen geographischen Region (area = Umgebung) zu benennen sind (Taylor/Braddock 2007). Zudem könnten die Daten einen erheblichen Selektionsbias aufweisen, weil weder die Befragtenpopulation noch die Rücklaufquote angegeben wurden. Die Herausgeber des *THES Ranking* behaupten hingegen, dass Reputationsumfragen, trotz ihrer Schwächen, die verlässlichsten Indikatoren für akademische Qualität liefern würden (Jobbins 2005).

Das Zahlenverhältnis zwischen Lehrenden und Studierenden kann schwer gemessen werden. Als Indikator für die Qualität der Lehre verwendet das *THES Ranking* das Zahlenverhältnis zwischen Lehrenden und Studierenden. Dies erscheint im Vergleich zu der im *ARWU Ranking* verwendeten Anzahl der Alumni, die einen Nobelpreis erhalten haben, als ein etwas angemessenerer Indikator für die Qualität der Studienprogramme. Jedoch hat er eine zweifelhafte Messvalidität, weil es, wie bereits oben erwähnt, nicht möglich ist, die Anzahl der Lehrenden und der Studierenden im internationalen Vergleich zuverlässig zu erheben. Die Anzahl der Zitationen pro akademischen Mitarbeiter ist ein adäquaterer Indikator für die Qualität der Forschung als die Anzahl der Publikationen insgesamt. Jedoch ergibt sich auch hier das Problem, die Anzahl der Forscher einer Universität zu bestimmen.

Ausländische Studierende und Forscher sind keine adäquaten Qualitätsindikatoren. Der im *THES Ranking* verwendete Indikator „Anteil der international mobilen Studierenden und Forscher“ ist ungeeignet, um die Qualität von Forschung oder Lehre im internationalen Vergleich darzustellen, weil dieser Anteil von vielen Faktoren abhängt, die nicht mit der Qualität in Verbindung stehen müssen. So spiegelt der Anteil der international mobilen Studierenden und ausländischen Forscher unter anderem die Ressourcen, die Reputation und die Unterrichtssprache wieder. Die Frage, ob Universitäten finanzielle Anreize haben, international mobile Studierende anzuziehen, oder ob Regularien bestehen, welche diese begrenzen, hat auch einen Einfluss auf den Internationalisierungsgrad von Universitäten. Aus den gleichen Gründen ist die Anzahl der ausländischen Forscher kein geeigneter Indikator für den internationalen Vergleich der Qualität in Forschung und Lehre.

Inadäquate Gewichtung der Indikatoren. Die Konstruktvalidität hängt auch davon ab, ob die Indikatoren so gewichtet sind, dass relevante Aspekte der Forschungs- und Lehrqualität eine adäquate Gewichtung erhalten. Das *THES Ranking* wird häufig dafür kritisiert, dass es Reputation und Internationalität mit 50 und 10% unverhältnismäßig viel Gewicht beimisst, während die Leistungen in der Forschung nur mit 20% einfließen (Taylor/Braddock 2007).

Das *THES Ranking* ist seit seiner Einführung starker Kritik ausgesetzt. Insbesondere problematisch ist, dass es eher ein Popularitätsmaß als eine Messung tatsächlicher Leistungen darstellt, weil es zu 50% Prozent auf den Meinungen von Forschern und Arbeitgebern basiert (Holmes 2006; Marginson 2007; Taylor/Braddock 2007). Daher hat es eine sehr zweifelhafte Konstruktvalidität. Das *ARWU* und das *Excellence Ranking* haben aufgrund ihrer objektiven Datenbasis und adäquateren Indikatoren von Leistungen in Forschung und Lehre eine höhere Konstruktvalidität (siehe nächstes Kapitel). Die Messvalidität des *THES Rankings* ist unter anderem aufgrund der niedrigen Rücklaufquoten gering. Einige Autoren erheben den Vorwurf, dass das *THES Ranking* gestaltet wurde, um britischen Universitäten einen Vorteil im internationalen Wettbewerb zu verschaffen. Viele britische Universitäten haben eine höhere Bekanntheit und einen höheren Anteil an ausländischen Studierenden als führende US-Universitäten, weisen aber bei Leistungsindikatoren oftmals schlechtere Werte auf (Marginson 2007; Holmes 2006). Marginson kommt daher zu dem Schluss, dass es sich bei dem *THES Ranking* um ein abgekartetes Spiel handelt: “The Times Higher ranking boosts the number of leading British universities and reduces the US universities in the world’s top 100 from 54 to 31 [A. K./H.-D. D.: compared to the *ARWU Ranking*] [...] this compounds the impression that the Times Higher rankings are a rigged game” [abgekartetes Spiel; A. K./H.-D. D.] (Marginson 2007, S. 134).

4.3 Das CHE Excellence Ranking

Messprobleme der Web of Knowledge-Literaturdatenbank. Unsere Literaturrecherche ergab, dass das *Excellence Ranking* des CHE bislang nur in einem wissenschaftlichen Artikel diskutiert wurde und dieser stammt von den Herausgebern des Rankings selbst (Berghoff/Brandenburg/Müller-Böling 2008). In diesem Ranking werden teilweise ähnliche Indikatoren wie im *ARWU Ranking* verwendet, weshalb ähnliche Stärken und Schwächen vorliegen. Ähnlich wie das *ARWU Ranking* verwendet das *Excellence Ranking* den Indikator „Anzahl der Publikationen der Vertreter des Fachs im *Web of Science* in den Jahren 1997-2004“. Die Probleme der Zuschreibung und der Bevorteilung großer Fachbereiche treffen daher auch hier zu. Auch wird der Indikator „Anzahl der *Highly Cited Researchers* in den Jahren 1984-2003“ verwendet. Dieser birgt auch hier das Problem, dass sich die Ergebnisse zum Teil auf Publikationen älteren Datums beziehen. Im Unterschied zum *ARWU Ranking* berücksichtigt das *Excellence Ranking* die Häufigkeit, mit der Publikationen zitiert werden. Das *Excellence Ranking* verwendet den von Van Raan entwickelten Indikator „Durchschnitt der pro Publikation erreichten Zitationen“ gewichtet mit der durchschnittlichen Zitierhäufigkeit des Fachgebiets, dem eine Publikation zugeordnet wird. Dieser sogenannte *crown indicator* ist besonders hilfreich, da er Universitäten belohnt, die konsistent Artikel mit einer überdurchschnittlichen Wirkung in

ihrem jeweiligen Fachgebiet publizieren. Der Indikator „Anzahl der Projekte im Marie-Curie-Förderprogramm“ lässt sich schwer einschätzen, da hierzu weitere Forschung darüber notwendig ist, ob sich aus jedem Land (respektive jeder Region) die jeweils besten Nachwuchswissenschaftlerinnen und -wissenschaftler bewerben und die besten Bewerber ausgewählt werden.

Höherer Informationsgehalt durch fach- und indikatorenspezifische Darstellung. Am *Excellence Ranking* fällt besonders positiv auf, dass die Werte der Universitäten nach den Indikatoren getrennt dargestellt werden und nicht versucht wird, die Qualität von Universitäten in einer Zahl zusammenzufassen. Auch ist es sinnvoll, dass einzelne Fachgebiete an Universitäten und nicht gesamte Universitäten die Analyseebene bilden. Die Literatur stimmt darin überein, dass eine solche Präsentation der Ergebnisse wünschenswert ist, weil weder Studierende noch Universitäten aus aggregierten Daten Schlüsse ziehen können (Usher/Savino 2006; Marginson/Van der Wende 2007). Insgesamt erhöhen die Wahl der Aussageebene (Graduiertenprogramm) und die Darstellung der Ergebnisse die Konstruktvalidität des *Excellence Rankings*. Einschränkend muss allerdings hinzugefügt werden, dass das *Excellence Ranking* vorgibt, ein Ranking von Graduiertenprogrammen zu sein. Es bleibt jedoch unklar, ob die vier verwendeten Indikatoren nur für Personen erhoben wurden, die im Graduiertenprogramm lehren, oder für das wissenschaftliche Personal der Universität respektive der Disziplin insgesamt.

5. Reliabilität der internationalen Rankings

Als weiteres Güterkriterium für internationale Rankings betrachten wir ihre Reliabilität. Sie bezieht sich auf die Frage, ob ein Messverfahren bei wiederholten Messungen die gleichen Ergebnisse produziert und ob die Ergebnisse mit anderen Qualitätssmessungen von Universitäten übereinstimmen. Das *THES* und das *ARWU Ranking* haben eine zweifelhafte Reliabilität, weil ihre Ergebnisse zeitlich stark schwanken. Im *THES Ranking* verändern sich die Ergebnisse für eine Universität von einem auf das andere Jahr nicht selten um hundert Rangplätze. Die *Emory University* stieg beispielsweise innerhalb eines Jahres von dem 173. auf den 56. Rang auf, während die *Purdue University* vom 59. auf den 127. Rang fiel (Marginson 2007). Es ist wenig plausibel, dass sich die Leistungen der Universitäten innerhalb kurzer Zeit derart dramatisch verändert haben. Die Reliabilität des *ARWU Rankings* wird darüber hinaus angezweifelt, weil die Ergebnisse nicht reproduziert werden konnten (Florian 2007). Weitere Zweifel an der Reliabilität der Rankings ergeben sich, weil Universitäten im *THES* und *ARWU Ranking* oftmals sehr unterschiedliche Rangplätze erhalten. Da beide Rankings vorgeben, die Qualität in Forschung und Lehre von Universitäten zu messen, müsste eine höhere Übereinstimmungsreliabilität resultieren. Das *Excellence Ranking* ist noch zu neu, um seine zeitliche Stabilität zu beurteilen.

6. Schlussfolgerungen für die Gestaltung von Rankings und Forschung über Rankings

Die Analyse internationaler Universitätsrankings bestätigt zwei zentrale Erkenntnisse über Rankings: Universitätsrankings spiegeln die Leistungen von Universitäten besser wieder, wenn sie die Leistungen (*Outcomes* und *Outputs*) in Forschung und Lehre betrachten und weniger die Prozesse oder *Inputs*. Dies ist der Fall, weil Letztere oftmals keine guten Prädiktoren für die Leistungen in Forschung und Lehre sind (Dill/Soo 2005). Outcome-Indikatoren (z. B. Zitationshäufigkeiten) und Output-Indikatoren (z. B. Publikationszahlen) sind für die Bewertung von Forschung besser geeignet als Input-Indikatoren (z. B. die Anzahl der Forscher). Die zweite Erkenntnis ist, dass Indikatoren nur verwendet werden sollten, wenn belegt ist, dass sie überhaupt einen Einfluss auf die *Outcomes* in Forschung oder Lehre haben können (Dill/Soo 2005). Indikatoren, wie die Anzahl der Alumni, die in ihrem späteren Leben einen Nobelpreis erhalten, haben gewiss keinen Einfluss auf die aktuelle Qualität der Lehre an ihrer früheren Universität.

Bibliometrische Analysen sind geeignet, um Forschungsleistungen international zu vergleichen. Zwar indiziert nicht jede Zitation einer Arbeit deren Qualität, aber sie zeigt in der Regel, dass die Publikation einen anerkannten Beitrag zum wissenschaftlichen Fortschritt oder zur wissenschaftlichen Debatte darstellt. Die Anzahl der Zitationen ist ein anerkannter Indikator für die wissenschaftliche Wirkung von Forschungsarbeiten (Bornmann/Daniel 2008). Jedoch spiegeln die Literaturdatenbanken *Web of Knowledge* und *SCOPUS*, auf denen viele bibliometrische Analysen basieren, die Forschungsleistungen in sozial- und geisteswissenschaftlichen Fächern nicht adäquat wider. Dies liegt unter anderem daran, dass Wissenschaftler dieser Fächer vor allem Bücher und in nicht-englischsprachigen Zeitschriften publizieren. Weitere Forschung ist nötig, um zu ermitteln, wie die Publikationsleistungen in den Geistes- und Sozialwissenschaften adäquat erfasst werden können. Zudem ist es wichtig, einheitliche Namen für die Universitäten zu verwenden sowie die Einträge in Literaturdatenbanken sorgfältiger durchzuführen, um ihre Zuverlässigkeit zu erhöhen.

Leistungen in der Lehre lassen sich wesentlich schwieriger als Forschungsleistungen international vergleichend beurteilen. Die im *ARWU* und *THES* Ranking verwendete Anzahl der Alumni, die einen Nobelpreis erhalten haben, und das Zahlenverhältnis zwischen akademischen Mitarbeitern und Studierenden sind als Indikatoren ungeeignet. Letzter Indikator kann nicht zuverlässig berechnet werden, weil, wie oben dargelegt, die Anzahl der akademischen Mitarbeiter und Studierenden kaum international vergleichend erhoben werden können. Zuverlässigere Informationen über die Qualität der Lehre können Befragungen von Studierenden und Absolventen liefern. Für das *Excellence Ranking* wurden zwar Studierende befragt, allerdings konnten die Ergebnisse aufgrund der geringen Rücklaufquote nur für einen Teil der Universitäten verwendet werden. Auch international vergleichende Tests von Studierendenleistungen, wie sie von der *OECD* geplant werden (*OECD* 2008), könnten im Prinzip wertvolle Informationen über die Qualität der Lehre bereitstellen. Allerdings wird es sehr schwierig sein, Kompetenztests zu entwickeln, welche das an Universitäten erworbene Wissen reliabel, valide und international vergleichbar messen (ACE 2000).

Neben den Problemen bei der Messung von Qualität in Forschung und Lehre, bedarf es weiterer Untersuchungen zu den Auswirkungen internationaler Rankings auf Universitäten, Studierende und Politik. Es wird befürchtet, dass Universitätsrankings die Universitäten dazu anreizen, ihre Lehrqualität oder ihr gesellschaftliches Engagement zu vernachlässigen, weil diese in Rankings nicht oder nur wenig adäquat erhoben werden (Dill/Soo 2005; Marginson/Van der Wende 2007). Auch wird prognostiziert, dass sie Studierende dazu anreizen, ihre Universitäten primär nach der Platzierung in Rankings auszusuchen oder der Politik Anreize geben, ihre Förderung auf wenige, exzellente Forschungsuniversitäten zu konzentrieren (Marginson/Van der Wende 2007). Die bisherige Forschung stellt fest, dass Universitätsrankings das Verhalten von Studierenden und Hochschulen beeinflussen, ist sich aber über die Intensität und Dauer des Einflusses uneinig (vgl. Daniel 2001; Ehrenberg 2002; Hazelkorn 2007; UNITE 2007; HEFCE 2008).

Es wird jedoch deutlich, dass Rankings insbesondere dann verzerrende Anreize ausüben, wenn sie die oben genannten Erkenntnisse nicht befolgen und Indikatoren verwenden, die *Inputs* oder Reputation messen und in einem geringen Zusammenhang mit Lehr- oder Forschungsqualität stehen (Dill/Soo 2005). Universitätsrankings, welche auf Outcomes, wie Zitationszahlen oder Absolventenerfolg basieren, geben den Universitäten und der Politik weniger verzerrende Anreize. Verzerrende Anreize können auch vermieden werden, indem vielfältige Aspekte der Leistungen von Hochschulen erhoben werden. In den *Berlin Principles on Ranking of Higher Education Institutions* (CHE/UNESCO/IHEP 2006) werden daher richtigerweise mehrdimensionale Rankings empfohlen, welche die Leistungen in der Lehre, der Weiterbildung oder dem Technologie-Transfer berücksichtigen, und damit eine adäquatere Beurteilung auch weniger forschungsintensiver Hochschulen erlauben würde.

Alle Hochschulrankings – auch solche, welche die genannten Verbesserungsvorschläge befolgen – können immer nur einzelne Aspekte von Hochschulqualität erfassen. Es besteht stets die Gefahr, dass Hochschulrankings falsch interpretiert werden und nicht als Informationen über einzelne Aspekte, sondern als globaler Qualitätsindikator verstanden werden. Daher ist es sehr wichtig, dass die Herausgeber von Hochschulrankings diese nicht als globale Bewertung von Hochschulen präsentieren, sondern darauf aufmerksam machen, welche Aspekte von Hochschulen sie mittels welcher Indikatoren bewerten und wo die Grenzen der Konstrukt- und Messvalidität liegen.

Auch die Medien und die Hochschulen tragen Verantwortung, Fehlinterpretationen von Hochschulrankings entgegen zu wirken. In der Presse und auf Hochschulwebseiten sollten nicht nur die Ergebnisse von Hochschulrankings, sondern immer auch die Grenzen der Indikatoren erwähnt werden.

Die erste Generation internationaler Hochschulrankings hat deutlich gemacht, dass weltweit ein großes Interesse an Leistungsvergleichen von Universitäten besteht. Die nächste Generation von Hochschulrankings steht allerdings vor der Herausforderung, der interessierten Öffentlichkeit auch valide und reliable Informationen über die Stärken und Schwächen von Universitäten zur Verfügung zu stellen.

Anmerkung

- 1 Die Literaturdatenbank *Web of Knowledge* von Thomson Reuters umfasst unter anderem die Datenbanken *Web of Science*, *Journal Citation Reports* und *Highly Cited Researchers*.

Literatur

- ACE 2008 = American Council on Education (2008): ACE Urges Caution Moving Forward with Proposed OECD International Student Learning Outcomes Proposal. URL: <http://www.acenet.edu> (Download: 22.09.2008).
- Aguillo, I. F./Ortega, J. L./Fernández, M. (2008): Webometric ranking of world universities: Introduction, methodology, and future developments. In: Higher Education in Europe, Vol. 33(2/3), pp. 233–244.
- Altbach, P. G. (2006): The dilemmas of rankings. In: International Higher Education, No. 42, pp. 1–2.
- Berghoff, S. /Brandenburg, U./Müller-Böling, D. (2008): Identifying the best: The CHE excellence ranking of European graduate programmes in the natural sciences and mathematics. In: Higher Education in Europe, Vol. 33(2/3), pp. 273–283.
- Bornmann, L./Daniel, H. D. (2008): What do citation counts measure? A review of studies on citing behavior. In: Journal of Documentation, Vol. 64(1), pp. 45–80.
- Braun, T./Szabadi-Peresztegi, Z./Kovács-Németh, É. (2003): No-bells for ambiguous lists of ranked Nobelists as science indicators of national merit in physics, chemistry and medicine, 1901–2001. In: Scientometrics, Vol. 56(1), pp. 3–28.
- Braun et al. 2007 = Braun, T./Dióspatonyi, I./Zádor, E./Zsindely, S. (2007): Journal gatekeepers indicator based top universities of the world, of Europe and of 29 countries. A pilot study. In: Scientometrics, Vol. 71(2), pp. 155–178.
- CHE 2007 = Center for Higher Education Development (2007): CHE Excellence Ranking. Identifying the Best: The CHE Ranking of Excellent European Graduate Programmes in the Natural Sciences and Mathematics. – Gütersloh.
- CHE/UNESCO/IHEP (2006): Berlin Principles on Ranking of Higher Education Institutions. URL: http://www.che.de/downloads/Berlin_Principles_IREG_534.pdf (Download: 22.09.2008).
- CWTS (2007): The Leiden Ranking. URL: <http://www.cwts.nl/cwts/LeidenRankingWebSite.html> (Download: 22.09.2008).
- Cybermetrics Lab (2007): Webometrics Ranking of World Universities. URL: <http://www.webofmetrics.info/> (Download: 22.09.2008).
- Daniel, H. D. (2001): Was bewirken Hochschul-Rankings? Wer orientiert sich an ihnen? In: Müller-Böling, D./Hornbostel, S./Berghoff, S. (Hrsg.): Hochschulranking. Aussagekraft, Methoden, Probleme. – Gütersloh, S. 121–124.
- Dill, D./Soo, M. (2005): Academic quality, league tables, and public policy: A cross-national analysis of university rankings. In: Higher Education, Vol. 49, pp. 495–533.
- EHrenberg, R. G. (2002): Reaching for the brass ring: The U.S. News & World Report rankings and competition. In: The Review of Higher Education, Vol. 26(2), pp. 145–162.
- Espeland, W. N./Sauder, M. (2007): Ranking and Reactivity: How Public Measures Recreate Social Worlds. In: American Journal of Sociology, Vol. 113(1), pp. 1–40.
- European Commission (2004): Mapping of Excellence in Economics. Office for Official Publications of the European Commission. – Luxembourg.
- Financial Times (2008): Financial Times Global MBA Ranking. URL: <http://rankings.ft.com/global-mba-rankings> (Download 22.09.2008).

- Florian, R. V. (2007): Irreproducibility of the results of the Shanghai academic ranking of world universities. In: *Scientometrics*, Vol. 72(1), pp. 25–32.
- Guarino et al. 2005 = Guarino, C./Ridgeway, G./Chun, M./Buddin, R. (2005): Latent variable analysis: A new approach to university ranking. In: *Higher Education in Europe*, Vol. 30(2), pp. 147–165.
- Hattendorf Westney, L. C. (2006): *Educational Rankings Annual 2006: Over 4600 Rankings and Lists on Education, Compiled from Educational and General Interest Published Sources*. – Gale Cengage.
- Hazelkorn, E. (2007): The Impact of League Tables and Ranking Systems on Higher Education Decision-Making. In: *Higher Education Management and Policy*, Vol. 19(2), pp. 81–05.
- HEFCE (2008): Counting What is Measured or Measuring What Counts? League Tables and Their Impact on Higher Education in England. Report to HEFCE by CHERI, Open University and Hobsons Research. Issues Paper 2008/14. – London.
- Holmes, R. (2006): The THES university rankings: Are they really world class? In: *Asian Journal of University Education*, Vol. 1, pp. 1–14.
- Jobbins, D. (2005): Moving to a global stage: a media view. In: *Higher Education in Europe*, Vol. 20(2), pp. 137–145.
- Liu, N. C./Cheng, Y. (2005a): The academic ranking of world universities. In: *Higher Education in Europe*, Vol. 30(2), pp. 217–228.
- Liu, N. C./Cheng, Y. (2005b): Academic ranking of world universities using scientometrics: A comment to the “Fatal Attraction”. In: *Scientometrics*, Vol. 64(1), pp. 101–109.
- Marginson, S. (2007): Global university rankings: Implications in general and for Australia. In: *Journal of Higher Education Policy and Management*. Vol. 29(2), pp. 131–142.
- Marginson, S./Van der Wende, M. (2007): To rank or to be ranked: The impact of global rankings in higher education. In: *Journal of Studies in International Education*, Vol. 11(3/4), pp. 306–329.
- Marsh, H. W./Hattie, J. (2002): The relation between research productivity and teaching effectiveness: Complementary, antagonistic or independent constructs? In: *Journal of Higher Education*, Vol. 73(5), pp. 603–633.
- Newsweek (2006): The world's most global universities, 13.08.06.
- OECD (2008): Proposed OECD Feasibility Study for the International Assessment of Higher Education Learning Outcomes (AHELO). URL: http://www.oecd.org/document/22/0,3343,fr_2649_37455_40624662_1_1_1_37455,00.html (Download: 22.09.2008).
- Shadish, W. R./Cook, T. D./Campbell, D. T. (2002): *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. – Boston.
- SJTU 2007 = Shanghai Jiao Tong University Institute of Higher Education (2007): Academic Ranking of World Universities 2007. URL: <http://ed.sjtu.edu.cn/ranking.htm>. (Download: 22.09.2008).
- SJTU 2008 = Shanghai Jiao Tong University Institute of Higher Education (2008): Academic Ranking of World Universities by Broad Subject Fields 2008. URL: <http://ed.sjtu.edu.cn/ARWU-FIELD2008.htm> (Download: 22.09.2008).
- Strathern, M. (Ed.) (2000): *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. – New York.
- Taylor, P./Braddock, R. (2007): International University Ranking Systems and the Idea of University Excellence. In: *Journal of Higher Education Policy and Management*. Vol. 29(3), pp. 245–260.
- THES 2007 = Times Higher Education Supplement (2007): Guide to the World's Top Universities. – London.
- Thomson ISI (2008): About ISI Highly Cited. URL: <http://www.isihighlycited.com> (Download: 22.09.2008).

UNITE (2007): The Student Experience Reports. – Bristol. URL: <http://www.unite-group.co.uk/Attachments/000171/Student%20Experience%20Report%202007.pdf> (Download: 5.7.2008).

Usher, A./Savino, M. (2006): A World of Difference: A Global Survey of University League Tables. Educational Policy Institute – Toronto.

Van Raan, A. F. J. (2005): Fatal Attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. In: *Scientometrics*, Vol. 62(1), pp. 133–143.