

## »DIE MAGIE DER ZAHLEN« – DIE HOCHSCHULVERRENKUNGEN VON AOL, MCKINSEY UND SPIEGEL

MAIA HÖDING, M.EIK MICHALKE, OLIVER NASS

für den Fachschafsrat Psychologie, Philipps-Universität Marburg  
Gutenbergstr. 18, 35037 Marburg – [fachs04@students.uni-marburg.de](mailto:fachs04@students.uni-marburg.de)

**ZUSAMMENFASSUNG.** Im November 2004 veröffentlichten die Unternehmensberatung McKinsey & Company, der Internetprovider AOL und das Politmagazin DER SPIEGEL ein neues Hochschulranking. Bei genauerer Betrachtung der präsentierten Ergebnisse konnte gezeigt werden, dass die Publikation erhebliche Mängel aufweist. Neben fehlender theoretischer Basis und einem Fragebogen, der weder testtheoretischen noch diagnostischen Grundvoraussetzungen entspricht, rücken auch methodisch unsinniges Vorgehen sowie offensichtlich falsche Ergebnisse simpelster Berechnungen die Veröffentlichung fern ab von dem, was man auch nur als »populärwissenschaftlich« bezeichnen könnte: Schon nach Korrektur des letzten Rechenschrittes verschieben sich 28 der 41 Hochschulen in dieser Veröffentlichung um bis zu sieben Rangplätze nach oben oder unten gegenüber denen, die ihnen DER SPIEGEL zuwies. Damit qualifiziert sich dieses Hochschulranking in der aktuellen Bildungsdebatte allenfalls als alarmierendes Negativbeispiel.

Das SPIEGEL-Ranking (Friedmann et al., 2004) ist Bestandteil einer Diskussion um die Reform der deutschen Hochschulen und stellt an sich den Anspruch, Argumente zu liefern, um in diese Diskussion steuernd einzugreifen. Veröffentlicht wurde eine Rangfolge von 41 Universitäten, an deren Spitze die besten und an deren Ende die schlechtesten Universitäten stünden. Laut SPIEGEL hat »*nie zuvor [...] eine Hochschulstudie umfassender die Qualifikationen und das Können der Studierenden in ganz Deutschland erfasst und analysiert*«, ihre Studie belege zugleich »*die wachsende Kluft zwischen studentischen Spitzenkräften und akademischem Mittelmaß*« (S. 178). Zentral für die Untersuchung ist der Begriff »Elite«, der jedoch nicht explizit definiert wird. Was mit diesem Begriff bezeichnet wird, kann nur indirekt geschlossen werden, indem die gemessenen Merkmale bzw. die zu ihrer Messung eingesetzten Fragen samt Antwortmöglichkeiten betrachtet werden. Bewertet werden sollen in der vorliegenden Umfrage generell nicht die Hochschulen selbst, sondern die an ihnen Studierenden.

Den AutorInnen dieser kritischen Auseinandersetzung mit dem veröffentlichten Ranking kamen (wie bei vorangegangenen Rankings anderer Magazine) sehr schnell Zweifel an den methodischen Kompetenzen der daran Beteiligten. Daher entschieden sie sich, diesem Ranking stellvertretend für weitere Veröffentlichung dieser Art etwas Zeit zu widmen und die gemachten Versprechungen auf ihre Haltbarkeit zu prüfen. Evaluation, diagnostische Verfahren und statistische Auswertungen sind schließlich ein zentraler Bestandteil der Diplom-Psychologie.

### OPERATIONALISIERUNG: DER FRAGEBOGEN

McKinsey & Company, AOL und DER SPIEGEL entwickelten einen Fragebogen (McKinsey & Company, 2004b), der online ausgefüllt werden sollte. In der empirisch-wissenschaftlich arbeitenden Psychologie ist es üblich, vor Erstellung eines Testinstrumentes theoretische Überlegungen mitzuteilen, die dem Messinstrument zugrundeliegen. Dies geschieht v. a. aus der Einsicht, dass vor dem Stellen einer Frage klar sein sollte, was man eigentlich in Erfahrung bringen

---

Wir danken Dr. Erwin Grüner und Prof. Dr. Hans-Henning Schulze, Arbeitsgruppe Methodenlehre und EDV, sowie Prof. Dr. Gerhard Stemmler, Arbeitsgruppe Differentielle und Diagnostische Psychologie, am Fachbereich Psychologie der Philipps-Universität Marburg, für ihren wissenschaftlichen Beistand beim Erstellen dieses Artikels.  
Dieses Dokument wurde mit freier Software erstellt – <http://www.lyx.org>.

möchte, um anschließend die Frage möglichst zielgerichtet stellen zu können.

Wie eingangs erwähnt, fehlt dem Ranking-Fragebogen eine öffentlich zugängliche, theoretische Basis. Die gestellten Fragen und ihre jeweiligen Antwortmöglichkeiten sind die einzigen Indizien, die auf das zugrundeliegende Bild der »Elite« schließen lassen, das erfasst werden soll. Über die Bildung der Fragengebiete wird jedoch bemerkt, dass daran maßgeblich »Auswertungen von Absolventenstudien, Arbeitgeberbefragungen und Stellenanzeigen« (Friedmann et al., 2004, S. 180) beteiligt waren.

Es wird erwähnt, dass gegebene Antworten mittels einer »detaillierten Metrik« zu einem Gesamtwert verrechnet wurden. Leider werden dazu jedoch ebenfalls keine genaueren Angaben gemacht: Gewichtung von Fragen und Antworten bleiben daher vollkommen im Unklaren.

Der Online-Fragebogen gliedert sich in eine Reihe thematischer Unterabschnitte. Generell werden die gegebenen Antworten in keiner Weise verifiziert: Wer angibt, im 17. Semester BWL an der TU München zu studieren und ein sechsmonatiges Praktikum in beobachtender Position bei ALDI-Süd gemacht zu haben, kann mit diesen Angaben in die Berechnung eingehen, selbst wenn in Wirklichkeit mangels Hochschulzugangsberechtigung ein Studium nicht einmal möglich wäre. Dies eröffnet selbstverständlich Manipulationsversuchen Tür und Tor. Auf welche Weise sich die Untersuchenden gegen derartige Manipulationen zu schützen suchten, ist im Abschnitt »Methode: Ausschluss von Datensätzen und Berechnung« auf Seite 4 zu erfahren.

Wir werden im Folgenden zunächst exemplarisch einige der Fragen<sup>1</sup> in der Reihenfolge ihres Auftretens genauer betrachten. Der vollständige Fragebogen kann in Anhang A auf Seite 12 eingesehen werden.

**Schul- und Studienleistungen.** In diesem frühen Abschnitt des Fragebogens werden bereits Werte erfragt, die auf »realistischen Einschätzungen« beruhen können (*»Mit welcher Note wurde Ihre Diplom- bzw. Magisterarbeit bewertet oder welche Note erwarten Sie?«*; *»Was war Ihr Notenschnitt im Examen oder was wird er nach Ihrer realistischen Einschätzung sein?«*).

Es werden also entgegen der im SPIEGEL vertretenen Behauptung nicht zwingend eindeutige Fakten, sondern mit hoher Wahrscheinlichkeit auch Wunschdenken der Teilnehmenden erfasst. Die fehlende Unterscheidung, ob die Angaben nun auf Tatsachen oder visionären Eingebungen beruhen, lassen den Eindruck zu, dass dies für die Untersuchenden keinen relevanten Unterschied macht.

**Studiendauer.** *»Mit wie vielen Fachsemestern insgesamt werden Sie voraussichtlich Ihr gesamtes Studium (Grund- und Hauptstudium) abschließen bzw. haben Sie Ihr Studium abgeschlossen?«* – Die Untersuchenden begleiten die ProbandInnen erneut in die Gefilde der Mutmaßungen – pikant. Eine zusätzliche Frage wie *»Wieviele Semester werden Sie voraussichtlich wegen unbesetzter Stellen oder nicht vorhandenen Seminarplätzen zusätzlich studieren müssen?«* wäre nicht unbedingt valider gewesen, hätte aber in diesem Zusammenhang ein zentrales Problem vieler Hochschulen treffend herausgehoben. Schließlich sind auch für realistisch hoch geschätzte Semesterzahlen nicht automatisch die Studierenden selbst verantwortlich.

**Sprachkenntnisse.** Die Angaben zu Sprachkenntnissen konnten in Form von Wahl einer der Kategorien »keine Kenntnisse«, »Grundkenntnisse«, »gute Kenntnisse«, »fließend in Wort und Schrift« und »Muttersprache« erfolgen. Auffällig ist hier die bewusste Eingrenzung auf »lebende« Sprachen (hauptsächlich Deutsch, Englisch, Französisch und Spanisch). Latein fehlt damit beispielsweise ebenso wie Alt-Griechisch vollständig. Die Kenntnisse in diesen Sprachen sind demnach für die anvisierte Bildungselite irrelevant.

**EDV-Kenntnisse.** Der Bereich der elektronischen Datenverarbeitung (EDV) beschränkt sich für die »besten« Studierenden an hiesigen Hochschulen auf Beispielanwendungen wie (ausdrücklich) Microsoft Office und »Programmiersprachen«. Als Antwortkategorien standen in beiden Fällen »Grundkenntnisse«, »breites Basiswissen«, »überdurchschnittliche Kenntnisse« und »Expertenwissen« zur Verfügung. Weder der Erfassungsraum der in Frage kommende Anwendungen, noch eine nähere Beschreibung des weiten Feldes der Programmiersprachen wird angegeben. Zugleich bleiben alle weiteren Gebiete

<sup>1</sup>in dieser Arbeit zunächst nicht weiter betrachtet haben wir die Bereiche »Berufserfahrung«, »Mobilität« und »Auszeichnungen«

der EDV unberücksichtigt (Netzwerk, Datenbanken, Instant Messaging, Groupware, Statistik-Software etc.). Da es sich bei der vorliegenden Untersuchung ausschließlich um einen online (am Computer) verfügbaren Fragebogen handelt, dürfte davon auszugehen sein, dass die Teilnehmenden in der Lage sind, mit einer Textverarbeitung einen Brief zu schreiben, so dass die erste Frage nicht viel Informationswert haben dürfte (man könnte ebenso schriftlich fragen, ob jemand lesen kann).

Unschärf bleibt allerdings vor allem, wie die Antworten auf den Kenntnisstand in »Programmiersprachen« interpretiert werden sollen. So glauben erfahrungsgemäß viele, dass die zum Erstellen von Homepages eingesetzte Auszeichnungssprache HTML eine Programmiersprache sei und könnten bei dieser Frage versucht sein, sich »überdurchschnittliche Kenntnisse« zuzuschreiben. Dagegen ist es für die produktive Arbeit mit tatsächlichen Programmiersprachen (etwa C, C++, Python, Java oder Mono) oft nur wichtig, ihre Grundprinzipien verinnerlicht zu haben und dies ggf. auf weitere Sprachen übertragen zu können. Viele ProgrammiererInnen beherrschen daher eine Reihe von Hochsprachen in Ansätzen, würden mit diesem »breiten Basiswissen« aber schlechter abschneiden als IT-Laien, die sich lediglich ihre eigene Homepage erstellen.

**Praxiserfahrung.** Unter »Praxiserfahrung« sind absolvierte Praktika von mindestens vier Wochen Dauer zu verstehen. Im Kontrast zu früheren Bereichen fragen die Untersuchenden an dieser Stelle nicht nach, wie viele dieser Praktika die TeilnehmerInnen nach »ihrer realistischer Einschätzung« bis zum Ende ihres Studiums noch absolvieren werden. Wer jedoch z. B. ein Marburger Psychologie-Diplom vorzuweisen hat, musste auch mindestens 28 Wochen Praktikum abgeleistet haben, welches viele erst spät im Studienverlauf machen.<sup>2</sup>

Im Gegensatz zu den angeblich EDV-Kenntnisse messenden »Programmiersprachen«, für die wie erwähnt nicht die Notwendigkeit gesehen wurde, auch nur ein einziges Beispiel zu nennen, finden die Teilnehmenden hier eine Auswahlliste von mehr als 250 Unternehmen vor. Interessante Details sind dabei in den »value«-Optionen des zugrundeliegenden HTML-Codes

zu finden. So steht zwischen A.T. Kearney (value=»1335«) und Accenture (value=»1336«) das Unternehmen ABB mit value=»1598« – die Liste der Auswahlmöglichkeiten wurde offenbar im nachhinein erweitert. Insgesamt wurden 29 Unternehmen nachträglich eingeordnet, allerdings kann nicht überprüft werden, zu welchem Zeitpunkt diese Erweiterung geschah, d. h. ob sich dies vollzog, während die Datenerhebung bereits lief.

Offensichtliches Berechnungskriterium ist die Anzahl der Mitarbeiter, wobei nicht genauer spezifiziert wird, ob hiermit das Personal einer lokalen Aussenstelle, des deutschen Tochterunternehmens oder des weltweit agierenden Konzerns gemeint ist, falls diese Unterscheidung für die eigene Praktikumsstelle relevant sein sollte. Gerade für den Fall einer offensichtlich nicht vollständigen Unternehmensliste kann es hierbei zu erheblichen Schwankungen der Angaben kommen.

**Engagement.** DER SPIEGEL wirbt wiederholt mit der Behauptung, endlich auch einmal das »soziale und politische Engagement« der Studierenden berücksichtigt zu haben. Werfen wir jedoch einen Blick auf die möglichen Auswahlkategorien, bleibt auch diese Aussage extrem fraglich, zumal nicht veröffentlicht wurde, welche Angaben welchen Beitrag zum abschließenden persönlichen Punktwert leisteten. Sind zwei Jahre Joggen (als »aktives Mitglied« engagiert im Bereich »Sport«) aus Sicht der Unternehmensberatung mehr wert als ein Jahr im Fachbereichsrat (»lokale Leitungsfunktion« unter »Hochschule/-gruppe, Schule«)? Und sollte sich die Technik-Referentin eines AStA, die Installationsparties zum Umstieg von Windows auf Linux organisiert, am besten unter »Hochschule/-gruppe, Schule«, »Politik«, »Soziales«, »Wissenschaft und Technik« oder im Zweifel neutral unter »anderes Engagement« einordnen?

• In welchem Bereich engagier(t)en Sie sich?

- Hochschule/-gruppe, Schule
- Kirche
- Kunst/Literatur
- Musik
- Politik
- Soziales
- Sport
- Umwelt

<sup>2</sup>und zwar nicht aus Schluderei, sondern um bei Kontakt zu »echten« Klienten möglichst viel psychologisches Vorwissen mitbringen zu können

- *Wirtschaft*
- *Wissenschaft und Technik*
- *Anderes Engagement*
- *Welche Rolle spiel(t)en Sie?*
  - *Passives Mitglied*
  - *Aktives Mitglied*
  - *Aufgabenübernahme*
  - *Lokale Leitungsfunktion*
  - *Übergeordnete Leitungsfunktion*

**Selbsteinschätzung.** In dieser Kategorie wird mit »Selbsteinschätzung« zum ersten mal eine fast zutreffende Beschreibung verwendet. Fragen nach der Einschätzung eigener »Leistungen und Qualifikationen im Vergleich zu allen Studenten, die in Deutschland im gleichen Fach und im gleichen Semester studieren« oder die Erwartungen, wo man sich »bei Studienabschluss im Vergleich anderer Absolventen im gleichen Fach in Deutschland« sehe, dürften jedoch allenfalls die Selbstüberschätzung der Teilnehmenden messen, nicht jedoch ihren tatsächlichen Stand im bundesweiten Vergleich.

Bemerkt werden sollte auch die Zusammenlegung von »öffentlichem Dienst« und »Non-Profit-Organisationen« als letzte Antwortalternative, wenn man im Anschluss an das Studium weder in »Wirtschaft und Industrie« noch in »Wissenschaft und Forschung« arbeiten möchte.

**Universität.** Im Gegensatz zu vergangenen Rankings sollte es in diesem explizit nicht um die Universitäten, sondern um die dort Studierenden gehen. Gefragt wird an dieser Stelle jedoch trotzdem, ob es an der eigenen Universität »spezifische Programme für den Besuch ausländischer Universitäten« gibt, »Kontakte zu potenziellen Arbeitgebern hergestellt« werden, »Graduiertenkollege« bestehen<sup>3</sup> und wie zufrieden man mit »der Qualität und den Studienbedingungen für Ihren Studiengang an Ihrer Universität« ist. Selbstverständlich wäre auch hier relevant zu erfahren, welche Auswirkungen diese Angaben auf das Gesamtergebnis haben.

Am Schluss folgt eine Einverständniserklärung und die Option, anderen von der Befragung zu berichten. Über diesen Service fällt es ggf. besonders leicht, Freunde und Bekannte zu bitten, in einem bestimmten Muster zu antworten, um die Ergebnisse beliebig zu manipulieren.

<sup>3</sup>wieviele Studierenden wissen dies überhaupt?

#### METHODE: AUSSCHLUSS VON DATENSÄTZEN UND BERECHNUNG

**Ausschlusskriterium IP-Adresse.** In den Angaben zur Methodik (McKinsey & Company, 2004a) wird beschrieben, dass auch die erfasste IP-Adresse herangezogen wurde, um über den Ausschluss von Datensätzen zu befinden. Dabei waren potentielle Ausschlusskandidaten Adressen, von denen mehrere Fragebögen vorlagen. Die offensichtliche Überlegung dahinter sind Manipulationsversuche durch mehrmalige Teilnahme. Technisch gesehen ist das gewählte Vorgehen jedoch aus zwei Gründen problematisch:

- (1) PC-Pools in Universitäten weisen den dort aufgestellten Rechnern i. d. R. immer die gleiche (statische) IP-Adresse zu. An den Rechnern arbeiten jedoch nacheinander verschiedene NutzerInnen. In diesen Fällen werden wahrscheinlich grundlos Datensätze von der Untersuchung ausgeschlossen.
- (2) Provider wie T-Online oder Arcor vergeben dagegen dynamische IP-Adressen, d. h. KundInnen dieser Unternehmen bekommen bei jeder Einwahl eine andere Adresse aus dem freien Pool zugewiesen. Wer es auf Manipulationen anlegt, braucht sich lediglich neu einzuwählen und entgeht der Entdeckung.

In die Studie drohen also durch den Einsatz dieser vermeintlichen Bereinigungsmethode Werte aufgenommen zu werden, die eigentlich ausgeschlossen werden sollten; andererseits werden recht wahrscheinlich Daten verworfen, die eigentlich berücksichtigt werden sollten. Im Rahmen der Veranstaltung »Statistik II«, zweites Semester Grundstudium in Diplom-Psychologie, werden diese Probleme gemeinhin als Fehler erster und zweiter Art bzw. als  $\alpha$ - und  $\beta$ -Fehler bezeichnet. Wie die Bezeichnung »Fehler« vermuten lässt, sind diese möglichst zu vermeiden, und nicht künstlich zu begünstigen.

**Ausschlusskriterium Fehlende Angaben.** Ebenfalls beschrieben wird ein Vorgehen zur Datensatzbereinigung, welches nach technischem Unverständnis auch statistisch-methodische Schwächen der Verantwortlichen vermuten lässt. Nachdem die nach eigenen Angaben etwa 80.000

Datensätze in Gruppen (keine Mängel, ein kleiner Mangel, zwei oder mehr kleine Mängel) eingeteilt wurden, sollte ein statistisches Verfahren zum Test von Mittelwertsunterschieden (t-Test) entscheiden, ob sich diese Gruppen »signifikant unterscheiden«. Es wurde ein signifikanter Unterschied gefunden und die Gruppe mit zwei oder mehr Mängeln aus der Untersuchung genommen.

Erstes Problem ist die Feststellung von »Mängeln«, die nach eigenen Angaben schon durch die Nicht-Überprüfbarkeit der angegebenen E-Mailadresse oder »zu schnelles« Ausfüllen des Fragebogens erreicht wurden. Schwieriger ist jedoch, dass Mittelwertsdifferenzen bei derart großen Stichproben nahezu immer signifikant ausfallen müssen. Bortz & Döring (2003) geben die optimale Stichprobengröße, um mit 99%iger Sicherheit ( $\alpha = 0.01$ ) auch kleine Effekte noch signifikant zu bekommen, mit 503 Personen an. Angaben zur relativeren Effekstärke (vgl. z. B. Cohen, 1969; Rosenthal, 1984) fehlen leider.

Die Untersuchenden erwähnen randständig, dass sie auch weitaus komplexere statistische Verfahren (Faktoren- und Regressionsanalysen) durchführten, was leider erneut mangels dargestellter Ergebnisse nicht beleuchtet werden kann.

**Berechnung der Ergebnisse.** Berücksichtigt wurden Fakultäten, für die mindestens 18 verwertbare Datensätze vorlagen. Die Gründe für diesen kritischen Wert von 18 Fällen bleiben im Unklaren, uns ist jedenfalls keine statistische Grundlage hierfür bekannt. Die mangelnde Sorgfältigkeit, die der Studie als Ganzes anlastet, lässt befürchten, dass dieser Wert gewählt wurde, um den Augenschein der Repräsentativität zu wahren (bei strengeren Kriterien wären logischerweise mehr Fachbereiche aus der Wertung gefallen). Ähnliche Fragen drängen sich auf, wenn als Kriterium für die Aufnahme in das Gesamtranking unbegründet die Zahl von mindestens acht gewerteten Fakultäten in den Raum gestellt wird. Ein Blick in die Tabelle zeigt, dass neun Hochschulen aus der Wertung fielen, wenn dieses Kriterium z. B. auf mindestens zehn Fachbereiche festgelegt worden wäre – unter ihnen die Spitzenposition der TU München.

<sup>4</sup>McKinsey teilte uns mit, man werde weder »anonymisierte Datensätze, noch Gewichtungen, Verteilungen oder vergleichende Analysen auf Basis von Einzelkriterien« veröffentlichen; dies sei »im Interesse der Teilnehmer«.

<sup>5</sup>Prof. Dr. Gerhard Arminger, Lehrstuhl für Wirtschaftsstatistik, Universität Wuppertal; Prof. Dr. Manfred Deistler, Institut für Wirtschaftsmathematik, Forschungsgruppe Ökonometrie und Systemtheorie, Technische Universität Wien; Der Beirat teilte uns mit, dem Ranking lediglich beratend zur Verfügung gestanden und die Berechnungen nicht selbst durchgeführt zu haben.

<sup>6</sup>OpenOffice Calc 1.1.3 – <http://www.openoffice.org>

## ERGEBNISSE

Die Ergebnisse sind nach unserem Dafürhalten vor allem eines: Sie sind falsch. Außerdem sind sie unvollständig; keine Angaben zur Geschlechter- oder Altersverteilung, keine Darstellung von Streuungen der Daten, nicht einmal eine genaue TeilnehmerInnenzahl (»knapp 50.000«), von fachbereichsspezifischen Angaben oder gar Vertrauensgrenzen ganz zu schweigen – die Preisgabe dieser deskriptiven Daten scheint den Untersuchenden nicht bedeutsam zu sein. Leider sind jedoch die abgedruckten Ergebnisse ohne stichprobencharakterisierende Angaben, wie sie in wissenschaftlichen Veröffentlichungen eine zwingende Selbstverständlichkeit darstellen, im Ganzen nicht seriös interpretierbar.

Wir bemühten uns, eine Kopie des Datensatzes von McKinsey zu erhalten, um die SPIEGEL-Ergebnisse auf dieser Basis nachzuvollziehen zu können bzw. die lückenhaften Angaben zu vervollständigen. Leider blieben diese Anfragen erfolglos,<sup>4</sup> so dass wir eine kritische Betrachtung der Ergebnisse nur aufgrund der dürftigen SPIEGEL-Veröffentlichung vornehmen können. Wie sich zeigt, lassen auch diese Zahlen bereits erhebliche Bedenken wachsen.

**Falsch berechnete Mittelwerte.** Wie uns sowohl der wissenschaftliche Beirat<sup>5</sup> der Untersuchung als auch McKinsey bestätigten, soll es sich bei den »Punktzahlen« in der Gesamtranking-Tabelle (Friedmann et al., 2004, S. 180) um einfache Mittelwerte handeln. Die zugrunde liegende Formel für die Berechnung des arithmetischen Mittels dürfte allgemein bekannt sein (vgl. z. B. Krauth, 1995):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Nach alarmierenden stichprobenartigen Taschenrechnerüberprüfungen übertrugen wir besagte Daten in ein Tabellenkalkulationsprogramm<sup>6</sup> und ließen die Mittelwerte erneut berechnen. Für 20 der 41 Werte erhielten wir auch nach mehreren Überprüfungen andere Ergebnisse, die sich nicht durch Rundungsfehler erklären ließen. Obwohl die Unterschiede von den

SPIEGEL-Angaben nur um 0.1-0.2 abweichen, ergäbe sich nach dem vom SPIEGEL vorgenommenen Ranking-Schema für 28 der gelisteten Universitäten eine z. T. deutlich andere Positionierung: Verschiebungen um bis zu sieben Rangplätze nach oben oder unten sind die Folge, wie Tabelle 1 auf Seite 8 zu entnehmen ist.

Unser Vorgehen stellt sich am Beispiel der Universität München wie folgt dar:

$$\begin{aligned}\bar{x}_{UM} &= \frac{3+1+3+3+3+3+3+3+3+3+2+3}{13} \\ &= 2.7692 \\ &\approx 2.8\end{aligned}$$

Laut SPIEGEL erhält die Universität München jedoch nur einen Punktwert von  $\bar{x} = 2.6$ . Theoretisch wäre es möglich, durch unterschiedliche Gewichtung der einzelnen Fakultäten einen vom arithmetischen Mittel abweichenden Wert zu erhalten. In diesem Beispielfall kommen jedoch nur genau die beiden Fachbereiche mit einem Punkt (Biologie) bzw. zwei Punkten (Soziologie) in Frage, um einen von  $\bar{x} = 3$  abweichenden Wert zu erklären. Es bedürfte einer interessanten Begründung, warum diese beiden Fachbereiche an dieser Universität derart stark gewichtet werden sollten, dass dies in einem Unterschied von 0.2 Punkten resultiert. Eine Gesamtübersicht der Mittelwertsergebnisse befindet sich in Anhang C auf Seite 16.

**Nicht vorhandene Unterschiede.** Ein grober Fehler, der populärwissenschaftlichen Statistiken regelmäßig unterläuft, ist die Beschränkung der Angaben auf Mittelwerte – selten wird die zugrundeliegende Stichprobengröße, und so gut wie nie die Streuung der Daten angegeben.<sup>7</sup>

So werfen die Tabellen Fragen auf, in denen für die einzelnen Studienfächer alle gewerteten Universitäten in einer Rangreihe gelistet werden. Die Einteilung in drei Leistungsgruppen (Spitzengruppe, Mittelfeld, Schlussgruppe, resultierend in drei, zwei oder einem Punkt für das Gesamtranking) wurde laut den Angaben zur Methodik (McKinsey & Company, 2004a) durch ein der Diskriminanzanalyse ähnliches Verfahren vorgenommen. Es sollte bei möglichst ähnlichen Gruppengrößen dafür sorgen, dass die Mittelwerte innerhalb einer Gruppe möglichst nah beieinander liegen, sich die Mittelwerte der Gruppen selbst jedoch möglichst stark voneinander unterscheiden. Besondere Aufmerksamkeit

wecken hier die teilweise extrem unterschiedlichen Streuungen der Mittelwerte, die bei der Einteilung offenbar nicht beachtet wurden. Da, wie erwähnt und schlechterdings üblich, keine Angaben zu den Streuungen der Einzelwerte veröffentlicht wurden, bleibt uns nur die weniger elegante Betrachtung der hochschulspezifischen Durchschnittswerte je Fachrichtung. Tabelle 2 auf Seite 9 stellt eine Übersicht über die Spannweiten dar, die wir aus den SPIEGEL-Tabellen errechneten.

Augenfällig wird dabei, dass z. B. (im unzulässigen Jargon des SPIEGELS gesprochen) im Fach Jura<sup>8</sup> die »beste« und die »schlechteste« Hochschule 22.9 Punkte auseinander liegen, wogegen die Differenz in Maschinenbau nur 5.9 Punkte beträgt. In Chemie liegt die »schlechteste« Hochschule der »Spitzengruppe« ganze 2.1 Punkte vor der »besten« in der »Schlussgruppe«, wogegen alleine die »Spitzengruppe« in BWL über 20.9 Punkte streut (von 30.0 insgesamt). Auch Statistiker wissen, dass Zahlen nicht alles sind, sondern sie im Kontext betrachtet werden müssen. Bei derart unausgewogenen Verteilungen muss sich die Frage ergeben, ob das gewählte Verfahren seinen Zweck erfüllt oder nicht doch ein anderes gewählt werden sollte. Den Untersuchenden schien daran gelegen, drei möglichst ausgeglichene Gruppen zu erhalten – dies könnte alternativ auch weniger kompliziert anhand der Gesamtstreuung geschehen. Das mag kein elaboriertes statistisches Verfahren sein (Kurzweiter, 2001), aber wenn wir uns erinnern, dass wir gar nicht wissen, was das eingesetzte Messinstrument überhaupt misst, dürfte dieses Vorgehen keinen weiteren Schaden verursachen. Es hätte zumindest den Vorteil, einzelne Hochschulen besser relativ zur gesamten Spannweite einzuordnen. Man hätte den Hochschulen für das Gesamtranking die Punktzahlen (3, 2 oder 1) also auch folgendermaßen zuweisen können: Man teile die Spannweite durch die Anzahl der gewünschten Gruppen und ziehe danach die Gruppengrenzen.

In BWL (Spannweite 30.0 Punkte) würde man eine gleichmäßige Gruppenspannweite von 10.0 Punkten erhalten, wonach die Schlussgruppe mit Mittelwerten bis 56.5, das Mittelfeld bis 65.5 und die Spitzengruppe bis zum Maximum von 76.5 reichen würde.

<sup>7</sup>die dahinterliegende Problematik dieser »Ergebnisse« können sich Nicht-Statistiker bei Interesse im Anhang B auf Seite 15 an einem kurzen Beispiel erläutern lassen

<sup>8</sup>genaugenommen müsste es Rechtswissenschaften heißen

	»exakt«	Durchschnitt			Ranking		
		rund	Spiegel	Differenz	real	Spiegel	Differenz
München, TU	3.0000	3	3	-	1	1	-
Freiburg, U	2.9167	2.9	2.9	-	2	2	-
Leipzig, U	2.8182	2.8	2.7	0.1	3	3	-
Berlin, HU	2.5385	2.5	2.6	-0.1	7	4	-3
Konstanz, U	2.6000	2.6	2.6	-	5	4	-1
München, U	2.7692	2.8	2.6	0.2	3	4	1
Heidelberg, U	2.5455	2.5	2.5	-	7	7	-
Stuttgart, U	2.6000	2.6	2.5	0.1	5	7	2
Tübingen, U	2.4167	2.4	2.5	-0.1	9	7	-2
Augsburg, U	2.3750	2.4	2.4	-	9	10	1
Mannheim, U	2.3750	2.4	2.4	-	9	10	1
Kaiserslautern, U	2.3750	2.4	2.3	0.1	9	12	3
Würzburg, U	2.2727	2.3	2.3	-	14	12	-2
Jena, U	2.3636	2.4	2.2	0.2	9	14	5
Münster, U	2.2308	2.2	2.2	-	15	14	-1
Berlin, FU	2.1000	2.1	2.1	-	16	16	-
Bonn, U	2.0000	2	2.1	-0.1	20	16	-4
Darmstadt, TU	1.9091	1.9	2.1	-0.2	23	16	-7
Dresden, TU	2.1333	2.1	2.1	-	16	16	-
Marburg, U	2.0000	2	2.1	-0.1	20	16	-4
Regensburg, U	2.0000	2	2.1	-0.1	20	16	-4
Saarbrücken, U	2.1250	2.1	2.1	-	16	16	-
Chemnitz, TU	2.1250	2.1	2	0.1	16	23	7
Erlangen-Nürnberg, U	1.9286	1.9	1.9	-	23	24	1
Göttingen, U	1.9167	1.9	1.8	0.1	23	25	2
Köln, U	1.9167	1.9	1.8	0.1	23	25	2
Mainz, U	1.6667	1.7	1.8	-0.1	28	25	-3
Trier, U	1.8889	1.9	1.8	0.1	23	25	2
Aachen, TH	1.6154	1.6	1.6	-	30	29	-1
Bremen, U	1.7000	1.7	1.6	0.1	28	29	1
Kiel, U	1.5833	1.6	1.6	-	30	29	-1
Berlin, TU	1.5455	1.5	1.5	-	32	32	-
Braunschweig, TU	1.5455	1.5	1.5	-	32	32	-
Frankfurt a.M., U	1.4615	1.5	1.5	-	32	32	-
Hamburg, U	1.5385	1.5	1.5	-	32	32	-
Bochum, U	1.4000	1.4	1.4	-	38	36	-2
Düsseldorf, U	1.3750	1.4	1.4	-	38	36	-2
Gießen, U	1.5000	1.5	1.4	0.1	32	36	4
Hannover, U	1.4545	1.5	1.4	0.1	32	36	4
Bielefeld, U	1.3000	1.3	1.3	-	40	40	-
Duisburg-Essen, U	1.1818	1.2	1.1	0.1	41	41	-
Abweichungen:				20			28

TABELLE 1. *Unterschiede der realen und der im Spiegel veröffentlichten Mittelwerte, sowie die sich daraus ergebenden Änderungen im Gesamtranking. Die »exakten« Werte dienen lediglich der Veranschaulichung, dass es sich hierbei nicht um Rundungsfehler handeln kann.*

Diese Einteilung hätte recht interessante Auswirkungen auf die vergebenen Punkte für das

Gesamtranking: Die ausgewiesenen Spitzenuniversitäten Mannheim, München (TU+U) und Stuttgart fielen somit aus der Spitzengruppe und

	Fachbereiche	Spannweiten				
		(N)	gesamt (Min/Max)	Schluss (Min/Max)	Mitte (Min/Max)	Spitze (Min/Max)
	<b>BWL</b>	58	30.0 (46.5/76.5)	5.7 (46.5/52.2)	2.9 (52.6/55.5)	20.9 (55.6/76.5)
	<b>Biologie</b>	43	11.2 (47.3/58.5)	4.0 (47.3/51.3)	1.8 (51.6/53.4)	4.9 (53.6/58.5)
	<b>Chemie</b>	43	8.7 (50.1/58.8)	4.4 (50.1/54.5)	1.3 (53.3/54.6)	2.2 (56.6/58.8)
	<b>Germanistik</b>	44	11.6 (49.8/61.4)	2.7 (49.8/52.5)	3.0 (53.0/56.0)	4.9 (56.5/61.4)
	<b>Politik</b>	43	15.1 (50.8/65.9)	4.6 (50.8/55.4)	3.9 (55.5/58.4)	7.4 (58.5/65.9)
	<b>Elektrotechnik</b>	26	11.7 (48.5/60.2)	5.0 (48.5/53.5)	2.3 (53.8/56.1)	3.5 (56.7/60.2)
	<b>Maschinenbau</b>	24	5.9 (51.9/57.8)	2.5 (51.9/54.4)	0.9 (54.5/55.4)	1.7 (56.1/57.8)
	<b>VWL</b>	23	11.2 (49.6/60.8)	2.7 (49.6/52.3)	1.7 (52.8/54.5)	5.0 (55.8/60.8)
	<b>Soziologie</b>	34	11.9 (47.9/59.8)	5.4 (47.9/53.3)	2.0 (53.5/55.5)	4.3 (55.5/59.8)
	<b>Jura</b>	41	22.9 (39.1/62.0)	5.0 (39.1/44.1)	1.9 (44.6/46.5)	15.2 (46.8/62.0)
	<b>Psychologie</b>	42	11.6 (52.5/64.1)	5.9 (52.5/58.4)	1.8 (58.5/60.3)	3.5 (60.6/64.1)
	<b>Medizin</b>	35	12.1 (48.0/60.1)	1.5 (48.0/49.5)	1.6 (49.6/51.2)	8.6 (51.5/60.1)
	<b>Informatik</b>	51	10.5 (50.1/60.6)	3.1 (50.1/54.2)	1.8 (54.6/56.4)	4.1 (56.5/60.6)
	<b>Mathematik</b>	33	11.8 (51.3/63.1)	5.0 (51.3/56.3)	2.9 (56.5/59.4)	3.5 (59.6/63.1)
	<b>Physik</b>	41	8.9 (53.9/62.8)	5.2 (53.9/59.1)	1.2 (59.7/60.9)	1.6 (61.2/62.8)

TABELLE 2. *Spannweiten der Mittelwerte je Fachbereich insgesamt und je Gruppenzuteilung (Schlussgruppe, Mittelfeld und Spitzengruppe). Die kursiven Werte stellen zusätzlich jeweils den niedrigsten und den höchsten Wert dar, der in den SPIEGEL-Tabellen angegeben ist.*

die Universität Leipzig sogar vom Mittelfeld in die Schlussgruppe. In Jura bliebe ausser einem Ausreisser<sup>9</sup> keine weitere »Spitzenuni« übrig, und auch hier fiel die Universität Leipzig gemeinsam mit Mannheim aus der Mitte in die Schlussgruppe. Ein ähnliches Bild ergäbe sich für Medizin: Das Spitzenfeld würde allein bestritten von der Privatuniversität Witten/Herdecke, u. a. würden München (TU+U), Freiburg und Heidelberg ins Mittelfeld sinken, und hier würde die Universität Leipzig sogar von ihrem letzten Platz im Spitzenfeld in die Schlussgruppe fallen.

Von den Einteilungsmodalitäten abgesehen, müsste zumindest auch gezeigt werden, dass diese Einteilung in X Gruppen gerechtfertigt ist, d. h. die Gruppen müssten sich im statistischen Sinne unterscheiden. Insbesondere bei nur wenigen Punkten auseinanderliegenden Werten scheint dies äußerst fraglich. Zuguterletzt müssen wir noch einmal darauf hinweisen, dass bei sämtlichen Durchschnittswerten nicht klar ist, was sie überhaupt messen. Selbst bei korrekter Rechnung ergäbe sich daher inhaltlich betrachtet kein »Ergebnis«.

**Top-Studenten.** Die Auszeichnung sogenannter »Top-Studenten« ist aufgrund des mangelhaften Fragebogens und der Tatsache, dass bereits 18

Teilnehmende einer Fakultät zur Berechnung genügen sollten, auf seriöse Weise unmöglich.

## DISKUSSION

Das vorliegende Ranking unterstreicht, wie wichtig eine kritische Überprüfung des derzeitigen Bildungssystems ist: Es ist einer theoretisch nicht ausreichend begründeten, methodisch unzureichend operationalisierten, am Ende sogar falsch gerechneten und unsachlich interpretierten Scheinstudie gelungen, das gesamte bundesdeutsche Hochschulwesen in Diskussion zu versetzen, ohne dass sich dieses auf die Weise Gehör verschafft hätte, die ihm eigentlich am nächsten liegen sollte: Wissenschaftliche Abgeklärtheit. Vor allem die Hochschulen, die schon kurz nach Erscheinen des SPIEGELS mit Überlegungen an die Öffentlichkeit gingen, als Konsequenz aus dem Ergebnis die Finanzierung ihrer Fachbereiche zu überdenken, müssen sich fragen lassen, ob sie ihre hohen Positionen zurecht einnehmen. Diese Frage wird drängender, nachdem der hiermit vorgelegte Mängelbericht aus dem Fachbereich Psychologie der Philipps-Universität Marburg stammt, der in diesem Ranking deutlich im Schlussfeld gewährt wurde.

<sup>9</sup>die Hamburger Bucerius Law School

Die unbegründete, von Beliebigkeit bestimmte Festlegung der Kriterien, was in die Wertung genommen wurde und was nicht, veranschaulicht die vollkommene Sinnfreiheit der vom SPIEGEL über 23 Seiten (Werbeanzeigen mitgezählt) verschiedentlich interpretierten Rankingtabelle: Hätte man nicht 18, sondern nur 15 oder mindestens 30 TeilnehmerInnen pro Fachbereich, oder statt der geforderten acht schon sechs oder erst zehn gewertete Fachbereiche als Kriterium für die Gesamtwertung herangezogen, würde die gesamte Rangreihe kräftig durchgewürfelt. Auch die Auswahl der überhaupt erfassten Studienfächer stellt eine Stellschraube der Beliebigkeit dar: Die Unterschiede im Fach Wirtschaftswissenschaften seien so groß gewesen, dass man dieses gänzlich nicht gewertet habe; Zahlen erfahren wir nicht. Ebenso unbelegt hätte man die Fächer Elektrotechnik oder Maschinenbau aus dem Gesamtranking herauslassen können, die schließlich nur 14 bzw. 13 der 41 Hochschulen im Programm haben.

Schon minimale Punktzahl-Unterschiede sorgten für Bewegung bei fast 75% der gelisteten Hochschulen. Mit diesen falsch berechneten Mittelwerten gerät nicht nur das tabellarische »Hauptargument« des SPIEGEL-Titels erheblich ins Wanken, sondern es stellt sich zudem die Frage nach den weiteren methodischen Qualitäten dieser Scheinstudie. Die fragwürdige Zuweisung der Punktwerte tut ihr Übriges – wie wir zeigten, hätte ein einfacheres und methodisch in keiner Weise weniger geeignetes Einteilungsverfahren zu erheblichen Punktverlusten der angeblichen Spitzenuniversitäten geführt. Statistik dreht sich oft um die Frage: »Kann das alles Zufall sein?« Man könnte auch fragen: Wurde hier nur in unverantwortlicher Weise gefuscht, oder gezielt auf ein Wunschergebnis hin schön gerechnet?

Der verwendete Fragebogen wird grundlegenden testtheoretischen Gütekriterien in keinster Weise gerecht. Ein Großteil der Fragen weist erhebliche Mängel betreffend seiner Validität<sup>10</sup> auf, was nicht zuletzt auf die völlig fehlende theoretische Basis zurückzuführen sein dürfte: Auf einem undefinierten »Elite«-Begriff kann auch nach Lektüre von Stellenanzeigen und Konsultationen mit Arbeitgebern nur schwer eine Überprüfung der Fragerichtung erfolgen. Genaugenommen wissen wir nun z. B. zwar, dass der Fachbereich Physik der Universität Bonn einen

Punktwert von 59.9 erreicht hat, und dass das 0.2 Punkte über dem der Technischen Universität Braunschweig lag. Wir wissen jedoch leider nicht, was diese Zahlen bedeuten. Zusätzlich darf allein aus den offensichtlichen Manipulationsmöglichkeiten bezweifelt werden, dass das Instrumentarium auch nur *ausreichende* Reliabilitätswerte<sup>11</sup> zu erreichen in der Lage ist.

Auch dieser SPIEGEL-Veröffentlichung gelingt es nicht, den Begriff der »Elite« klar zu definieren. Bei fehlender Definition des Begriffes ist eine Diskussion über Für und Wider jedoch relativ sinnfrei. So bleibt in diesem Fall »Elite« ein Euphemismus für wirtschaftlich abverlangtes Spitzen-Humankapital, dass vor allem durch Stärken in Selbstdarstellung und Selbstbewusstsein gute Erfolgsaussichten hat, während die tatsächlichen »Qualitäten« unergründet bleiben. Laut Deutschland-Chef von McKinsey, Jürgen Kluge, sei der »Studentenspiegel« das erste »fundierte« Hochschulranking, das sich an den Anforderungen des Arbeitsmarktes orientiert (Friedmann et al., 2004, S. 184). Fragen und Antwortmöglichkeiten sind demnach einseitig auf die Bedürfnisse von Firmen und Unternehmen zugeschnitten, wenn auch oft sprachlich getarnt: Die im Management so wichtigen »Soft skills« versucht der SPIEGEL als »soziales Engagement« zu verkaufen. Die proklamierte Fundiertheit lässt dagegen wie gezeigt in allen Punkten zu wünschen übrig.

Es herrscht allgemeiner Konsens darüber, dass die »Elite« von morgen schneller ausgebildet sein müsse. So findet an den Hochschulen bereits ein großer Umbau zu verschulden Bachelor/Master-Studiengängen statt. Die falschen Mittelwerte in der Rankingtabelle bemerkten in Marburg jedoch zwei sogenannte »Langzeitstudenten« mit einem Taschenrechner, nachdem sie im Rahmen ihres Diplomstudiengangs selbstbestimmt lernen konnten, den eigenen Kopf zu benutzen. Vor diesem Hintergrund gewinnt zumindest die schon vom SPIEGEL festgestellte »*wachsende Kluft zwischen studentischen Spitzenkräften und akademischem Mittelmaß*« eine neue Qualität.

Auch die zahlreichen KritikerInnen von Rankings wie diesem müssen sich kritisieren lassen: Vielfach beschränken sie sich in ihren Entgegnungen und Abweisungen darauf, auf den

<sup>10</sup>Validität: »Misst ein Test überhaupt, was er messen soll?«

<sup>11</sup>Reliabilität: »Wie zuverlässig misst ein Test, was auch immer er misst?«

selbst empfundenen, zugrundeliegenden »Elite-Gedanken« mit Gegenargumenten zu antworten. Dabei bleibt von eben diesem bei genauerer Betrachtung wie in diesem Fall meist nicht viel übrig. Vielmehr laden die methodischen Unzulänglichkeiten zu Kritik ein, die letztendlich auch wieder auf die fehlende inhaltliche Basis zurückfallen. Berechtigte Fragen wären unseres Erachtens zum Beispiel: Möchten sich Politik und Gesellschaft überhaupt Rat holen in Fragen der zukünftigen Bildungsstruktur, von einem Politikmagazin, einem Internetprovider und einer Unternehmensberatung, denen es auch in Kooperation mit Wirtschaftsstatistikern nicht gelingt, einen einfachen Mittelwert zu berechnen? Wird die Zukunft unserer Hochschullandschaft mitentschieden von mangelhaften Minderleistungen, die nicht annähernd an das Niveau einer Semesterarbeit im diplompsychologischen Grundstudium heranreichen? Sieht die Wissenschaft von morgen so aus, dass große Namen mit Gütekriterien verwechselt werden? Welche eigentlichen Interessen verfolgen die Fürsprecher einer unscharf gehaltenen »Elite«, wenn sie ihre Argumentationen mit scheinwissenschaftlichen Publikationen wie solchen Rankings untermauern müssen? Warum sollen Investitionen zukünftig weniger in die Breite als mehr in eine »Spitze« gehen, die zudem vielleicht überhaupt nicht diese »Spitze« darstellt?

Abschließend müssen wir befinden, dass mit der Anschaffung eines kostengünstigen Taschenrechners vielfach ein gewichtigerer Beitrag zur Diskussion um erforderliche Reformationen des Bildungssystems geleistet werden kann als durch ein von atheoretischem Wunschenken getriebenes, unwissenschaftlich operationalisiertes und letztendlich falsch gerechnetes Ranking.

Kurzweiser, H. (2001). Faking quotes, ignoring facts and folding the truth: Testing the carelessness of subjects reading scientific articles. *Journal of Statistics, Accuracy and Lies*, 23, 42, 665–667.

McKinsey & Company (2004a). Methodik des Studentenspiegels. WWW: <http://www.studentenspiegel.de/methodik.pdf>; 9.12.2004.

McKinsey & Company (2004b). Studentenspiegel. WWW: <http://www.studentenspiegel.de/>; 9.12.2004.

Rosenthal, R. (1984). *Meta-analytic procedures for social research* Beverly Hills, CA: Sage.

#### LITERATUR

Bortz, J. & Döring, N. (2003). *Forschungsmethoden und Evaluation* Berlin: Springer.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences* New York: Academic Press.

Friedmann, J., Hackenbroch, V., Hipp, D., Klawitter, N., Koch, J., Lakotta, B., Mohr, J., Schmitz, C., Thimm, K. & Wüst, C. (2004). Die Elite von Morgen – Wo studieren die Besten? *DER SPIEGEL*, 48, 178–200.

Krauth, J. (1995). *Testkonstruktion und Testtheorie* Weinheim: Psychologie Verlags Union.

## ANHANG A. DER FRAGEBOGEN

- **STARTSEITE**
  - Bitte teilen Sie uns zunächst mit, was Sie studieren bzw. studiert haben
  - An welcher deutschen Universität studier(t)en Sie?
  - Was ist/war Ihr Studien-Hauptfach?
  - In welchem Studienabschnitt befinden Sie sich?
  - Mit welchem Abschluss endet(e) Ihr Studium?
  - In welchem Fachsemester befinden Sie sich zur Zeit (im momentan angestrebten Abschluss)?
  - Studieren Sie einen zweiten Studiengang (Doppelstudium) oder haben Sie bereits ein anderes Studium abgeschlossen?
- **SCHUL- UND STUDIENLEISTUNGEN**
  - Welchen Notendurchschnitt hatten Sie im Abitur?
  - Welchen Notendurchschnitt hatten Sie im Vordiplom bzw. im Physikum oder der Zwischenprüfung?
  - Was ist Ihr aktueller Notendurchschnitt (nur examensrelevante Noten aus dem Hauptstudium)?
  - Mit welcher Note wurde Ihre Diplom- bzw. Magisterarbeit bewertet oder welche Note erwarten Sie?
  - Was war Ihr Notenschnitt im Examen oder was wird er nach Ihrer realistischen Einschätzung sein?
- **STUDIENDAUER**
  - Mit wie vielen Fachsemestern haben Sie Ihr Grundstudium abgeschlossen oder werden Sie es voraussichtlich abschließen?
  - Mit wie vielen Fachsemestern insgesamt werden Sie voraussichtlich Ihr gesamtes Studium (Grund- und Hauptstudium) abschließen bzw. haben Sie Ihr Studium abgeschlossen?
  - Wie alt sind Sie?
- **SPRACHKENNTNISSE**
  - Wie gut sind Ihre Kenntnisse in der jeweiligen Sprache?
    - Deutsch
    - Englisch
    - Französisch
    - Spanisch
  - Haben Sie Kenntnisse in weiteren lebenden Sprachen?
- **SPRACHKENNTNISSE (WEITERE SPRACHEN)**
  - Geben Sie hier bitte an, in welchen weiteren Sprachen Sie Kenntnisse haben und wie gut diese Kenntnisse sind:
    - Weitere Sprache 1 -3
    - Kenntnisse weitere Sprache 1-3
- **EDV-KENNTNISSE**
  - Welche Kenntnisse besitzen Sie in den folgenden Bereichen der EDV?
    - Anwenderkenntnisse (z. B. Microsoft Office)
    - Programmiersprachen
- **BERUFSERFHRUNG**
  - Haben Sie Zivildienst, Wehrdienst oder ein Freiwilliges Soziales Jahr geleistet? (nein/ja)
  - Hatten Sie vor oder während des Studiums Berufserfahrung (keine Praktika)? Bitte gehen Sie auf die beiden wichtigsten Tätigkeiten näher ein.
    - Art der 1.-2. Berufserfahrung
    - Dauer der 1.-2. Berufserfahrung
  - Waren oder sind Sie während Ihres Studiums als studentische Hilfskraft tätig (mit mind. 30 Stunden Arbeitszeit/Monat)?
- **PRAXISERFAHRUNG**

- Wie viele Praktika mit mindestens 4 Wochen Dauer haben Sie vor oder während Ihres Studiums außerhalb der Universität absolviert?
- **PRAXISERFAHRUNG PRAKTIKUM 1**
  - Bitte beschreiben Sie Ihre Praktika näher. Sollten Sie mehr als 3 Praktika absolviert haben, führe sie bitte die 3 Wichtigsten auf.
  - Bitte nennen Sie Zeitraum und inhaltliche Aufgaben des Praktikums.
    - Beginn (Jahr)
    - Dauer (Monate)
    - Inhaltliche Aufgaben
  - Wo haben Sie das Praktikum absolviert?  
Falls die Liste das Unternehmen/die Organisation nicht enthält, nennen Sie bitte Namen und Mitarbeiteranzahl.
    - Name des Unternehmens
- **MOBILITÄT**
  - In wie vielen verschiedenen Städten haben Sie seit Studienbeginn jeweils mindestens 6 Wochen gewohnt?
  - Wie viel Zeit haben Sie für Studium, Praktika, Sprachkurse, etc. im Ausland verbracht (ohne Urlaubsreisen)?
- **ENGAGEMENT**
  - Haben Sie sich neben dem Studium, der Schule oder der Ausbildung engagiert? Wie viele solcher Aktivitäten haben Sie regelmäßig betrieben?
- **ENGAGEMENT 1**
  - Auf den folgenden Seiten beschreiben Sie bitte Ihr Engagement näher. Sollten Sie mehr als 3 Aktivitäten angegeben haben, konzentrieren Sie sich bitte auch die drei Wichtigsten.
  - In welchem Bereich engagier(t)en Sie sich?
    - Hochschule/-gruppe, Schule
    - Kirche
    - Kunst/Literatur
    - Musik
    - Politik
    - Soziales
    - Sport
    - Umwelt
    - Wirtschaft
    - Wissenschaft und Technik
    - Anderes Engagement
  - Über welchen Zeitraum engagier(t)en Sie Sich regelmäßig?
    - unter 6 Monate
    - 6 bis unter 12 Monate
    - 1 bis unter 2 Jahre
    - 2 bis unter 3 Jahre
    - 3 Jahre oder mehr
  - Welche Rolle spiel(t)en Sie?
    - Passives Mitglied
    - Aktives Mitglied
    - Aufgabenübernahme
    - Lokale Leitungsfunktion
    - Übergeordnete Leitungsfunktion
- **AUSZEICHNUNGEN, VERÖFFENTLICHUNGEN**
  - Haben Sie während Ihres Studiums Stipendien erhalten oder erhalten Sie diese? Bei mehr als zwei Stipendien geben Sie bitte die beiden wichtigsten an
    - Stipendium 1-2

- Haben Sie während Ihres Studiums oder davor Preise für besondere Leistungen erhalten?
- Haben Sie in einer Sportart Platzierung(en) oder einen Preis (Preise) errungen; wenn ja auf welchem Niveau?
- Haben Sie während Ihres Studiums (oder davor) wissenschaftliche Arbeiten veröffentlicht? (Nein Ja)
- SELBSTEINSCHÄTZUNG
  - Wie schätzen Sie Ihre Leistungen und Qualifikationen im Vergleich zu allen Studenten ein, die in Deutschland im gleichen Fach und im gleichen Semester studieren?
  - Und am Ende des Studiums? Wo erwarten Sie bei Studienabschluss im Vergleich anderer Absolventen im gleichen Fach in Deutschland zu stehen?
  - In welchem Bereich möchten Sie nach Abschluss des Studiums bzw. der Promotion arbeiten?
    - Weiß noch nicht
    - Wirtschaft und Industrie
    - Öffentl. Dienst und Non-Profit-Organisationen
    - Wissenschaft und Forschung
  - Wann werden Sie Ihr Studium voraussichtlich beenden bzw. wann haben Sie es beendet?
    - Monat
    - Jahr
- UNIVERSITÄT
  - Gibt es an Ihrer Universität für Ihr Studienfach spezifische Programme für den Besuch ausländischer Universitäten?
  - Werden an Ihrer Universität für Ihr Fach Kontakte zu potenziellen Arbeitgebern hergestellt?
  - Bestehen an Ihrer Universität für Ihr Fach Graduiertenkollege?
  - Wie zufrieden sind Sie mit der Qualität und den Studienbedingungen für Ihren Studiengang an Ihrer Universität
- EINVERSTÄNDNISERKLÄRUNG
  - Damit wir Ihnen Ihr individuelles Qualifikationsprofil zusenden und Sie an unserem Gewinnspiel teilnehmen können, bitten wir Sie im Folgenden um die Eingabe Ihrer personenbezogenen Daten. Wir versichern Ihnen, dass wir diese streng vertraulich behandeln werden (Hinweise zum Datenschutz).
  - Um Ihre Angaben speichern und verarbeiten können, bedarf es Ihrer ausdrücklichen Zustimmung. Weiterhin bieten wir Ihnen an, Sie auch in Zukunft per E-Mail über Aktuelles zu den Themen "Bildung" und "Karriere " zu informieren, so auch über eine mögliche Folgebefragung. Auch für diesen Service werden wir Ihre Daten selbstverständlich nicht an Dritte weitergeben.
  - Ich bin damit einverstanden, dass meine personenbezogenen Daten gespeichert werden, um mein individuelles Qualifikationsprofil an mich zu senden und meine Teilnahme am Gewinnspiel zu ermöglichen. (nein/ja)
- TELL-YOUR-FRIEND
  - Vielen Dank für Ihre Teilnahme!  
Wir haben Ihre Antworten gespeichert und werden Ihnen Ihr vorläufiges Profil zusenden, sobald eine hinreichend große Zahl von Studenten Ihres Fachs an der Befragung teilgenommen hat. Ihr endgültiges Profil erhalten Sie nach Abschluss der Befragung. Zu diesem Zeitpunkt werden wir auch die Verlosung vornehmen. Wenn Ihnen die Befragung gefallen hat und Sie Ihre Kommilitonen darüber informieren wollen, können Sie im Folgenden deren E-Mail-Adresse angeben. Damit erhöhen Sie gleichzeitig Ihre Chancen auf einen Gewinn bei unserem Gewinnspiel: Für jeden empfohlenen Teilnehmer, der den Fragebogen vollständig ausfüllt, erhalten Sie ein weiteres Los, und können so bis zu sechs Lose sammeln. Ich möchte den Studentenspiegel an folgende Personen weiterempfehlen (bitte E-Mail-Adresse angeben):

## ANHANG B. ERLÄUTERUNGEN ZUR BEDEUTUNG VON STREUUNGEN

Nehmen wir an, wir hätten zwei Datensätze mit je drei Messwerten (9, 10 & 11 bzw. 1, 2 & 27). Während der erste Datensatz bei einer Spannweite von 2 Punkten recht gering streut, weist der zweite Datensatz bei einer Spannweite von 26 Punkten eine deutlich höhere Streuung auf. Beide Datensätze haben jedoch den gleichen Mittelwert von 10. Während dieser Mittelwert für den ersten Datensatz eine recht gute Beschreibung darstellt (alle Werte liegen relativ nah an ihm), charakterisiert er den zweiten Datensatz nur sehr ungenau.

Betrachten wir zusätzlich einen dritten Datensatz (5, 7 & 24), der damit einen Mittelwert von 12 hat, so täuscht dieser höhere Mittelwert darüber hinweg, dass zwei der drei Werte kleiner sind als der Vergleichsmittelwert 10 und zudem der größte Wert (24) immer noch kleiner ist als der größte Wert des zweiten Datensatzes (27) – die Verteilungen der Einzelwerte überschneiden sich derart stark, dass es statistisch gesehen nicht gerechtfertigt wäre, einen »signifikanten« Unterschied zwischen den Datensätzen festzustellen, d.h. die Differenz zwischen 10 und 12 müsste auf zufällige Schwankungen zurückgeführt werden, wenn man die Daten seriös interpretiert.

## ANHANG C. GESAMTTABELLE MITTELWERTE

	BWL	Bio	Chemie	Elektrotechnik	Germanistik	Informatik	Jura	Maschinenbau	Mathe	Medizin	Physik	Politologie	Psychologie	Soziologie	VWL	Summe	Durchschnitt (exakt)	Durchschnitt (rund)	Durchschnitt (Spiegel)	Differenz Durchschnitt	Rang (real)	Rang (Spiegel)	Differenz Ranking
München, TU	3	3	3	3		3		3	3	3	3				27	3.0000	3	3	-	1	1	0	
Freiburg, U		3	3		3	3	3		3	3	3	3	3	3	2	35	2.9167	2.9	2.9	-	2	2	0
Leipzig, U	2	3	3		3	3	2			3	3	3	3	3		31	2.8182	2.8	2.7	0.1	3	3	0
Berlin, HU	2	3	1		3	2	3		2	3	3	3	2	3	3	33	2.5385	2.5	2.6	-0.1	7	4	-3
Konstanz, U		3	3		2	3	3				2	3	2	2	3	26	2.6000	2.6	2.6	-	5	4	-1
München, U	3	1	3		3	3	3		3	3	3	3	3	2	3	36	2.7692	2.8	2.6	0.2	3	4	1
Heidelberg, U		3	3		3		3		3	3	2	3	2	1	2	28	2.5455	2.5	2.5	-	7	7	0
Stuttgart, U	3	3	2	3	1	3		3	3		3	2				26	2.6000	2.6	2.5	0.1	5	7	2
Tübingen, U	2	3	3		3	2	3			3	3	2	1	1	3	29	2.4167	2.4	2.5	-0.1	9	7	-2
Augsburg, U	2				3	3	2		3		3	1		2		19	2.3750	2.4	2.4	-	9	10	1
Mannheim, U	3				1	3	2					2	3	2	3	19	2.3750	2.4	2.4	-	9	10	1
Kaiserslautern, U	3	1	2	3		2		3	3		2					19	2.3750	2.4	2.3	0.1	9	12	3
Würzburg, U	2	2	3		3	2	2			3	2	2	2		2	25	2.2727	2.3	2.3	-	14	12	-2
Jena, U	3	2	3		3	3	2			2	3	2	2	1		26	2.3636	2.4	2.2	0.2	9	14	5
Münster, U	3	1	2		2	3	2		2	3	1	2	3	3	2	29	2.2308	2.2	2.2	-	15	14	-1
Berlin, FU	1	2	3		2	2	2				3	3	2		1	21	2.1000	2.1	2.1	-	16	16	0
Bonn, U		1	1		2	1	3		2	3	2	2	2	3	2	24	2.0000	2	2.1	-0.1	20	16	-4
Darmstadt, TU		1	2		1	3		3	2		3	3	2	1		21	1.9091	1.9	2.1	-0.2	23	16	-7
Dresden, TU	2	3	2	2	2	2	1	2	2	1	2	2	3	3	3	32	2.1333	2.1	2.1	-	16	16	0
Marburg, U	2	2	3		2	2	1			2	3	1	1	3		22	2.0000	2	2.1	-0.1	20	16	-4
Regensburg, U	2	3	2		2		2		2	2	2	1	3		1	22	2.0000	2	2.1	-0.1	20	16	-4
Saarbrücken, U	1		2		3	3	3			2	2		1			17	2.1250	2.1	2.1	-	16	16	0
Chemnitz, TU	2			3	2	3		1				2	3	1		17	2.1250	2.1	2	0.1	16	23	7
Erlangen-Nürnberg, U	3	2	1	2	3	2	1	2	3	1	2	1	2	2		27	1.9286	1.9	1.9	-	23	24	1
Göttingen, U	2	1	1		3		2		3	1	2	2	2	2	2	23	1.9167	1.9	1.8	0.1	23	25	2
Köln, U	2	2	1		1		2		2	2	1	2	2	3	3	23	1.9167	1.9	1.8	0.1	23	25	2
Mainz, U	2	1	2		2		2		1	1	2	2	1	3	1	20	1.6667	1.7	1.8	-0.1	28	25	-3
Trier, U	2				2	1	2		3			2	1	2	2	17	1.8889	1.9	1.8	0.1	23	25	2
Aachen, TH	1	2	1	3	2	2		1	2	1	1	1	3	1		21	1.6154	1.6	1.6	-	30	29	-1
Bremen, U	2	2		1	3	1	1				2	2	1	2		17	1.7000	1.7	1.6	0.1	28	29	1
Kiel, U	2	2	1	2	1	2	2			2	2	1	1		1	19	1.5833	1.6	1.6	-	30	29	-1
Berlin, TU	2		1	2	1	2		1	2		1		1	3	1	17	1.5455	1.5	1.5	-	32	32	0
Braunschweig, TU	2	2	1	2	2	1		2	1		2	1		1		17	1.5455	1.5	1.5	-	32	32	0
Frankfurt aM, U	1	1	2		1	1	1		2	2	3	1	1	1	2	19	1.4615	1.5	1.5	-	32	32	0
Hamburg, U	1	2	1		3	2	1		1	1	1	3	1	2	1	20	1.5385	1.5	1.5	-	32	32	0
Bochum, U		1	1	2	2		1	1	2	2	1		1			14	1.4000	1.4	1.4	-	38	36	-2
Düsseldorf, U	1	1	1		2		2			2			1	1		11	1.3750	1.4	1.4	-	38	36	-2
Gießen, U	1	1			2		1			2	1	3	1			12	1.5000	1.5	1.4	0.1	32	36	4
Hannover, U		3	1	1	1	3	1	1	2		1	1		1		16	1.4545	1.5	1.4	0.1	32	36	4
Bielefeld, U	1	1	1		1	2	1		1		1		1	3		13	1.3000	1.3	1.3	-	40	40	0
Duisburg-Essen, U	1		1	1	1	3		1	1	1	1	1		1		13	1.1818	1.2	1.1	0.1	3	-	0

TABELLE 3. Punkte und Mittelwerte