

eValuation – Neue Medien in der Lehrevaluation

Philipp Pohlenz, Jan-Peter Hagenmüller

Rektorat, Servicestelle für Lehrevaluation
Universität Potsdam
Am Neuen Palais 10
14469 Potsdam
pohlenz@rz.uni-potsdam.de
jphagen@rz.uni-potsdam.de

Abstract: Lehrevaluation in Form von Erhebungen studentischer Qualitätsurteile zu Lehre und Studium kann als Anwendungsfall der empirischen Sozialforschung aufgefasst werden. Die Abwicklung der Datenerhebungen mittels des Internet stellt dabei eine kostengünstige und einfach zu handhabende Alternative zu herkömmlichen „paper&pencil Verfahren“ dar. Der Einsatz neuer Medien in der Lehre trägt in diesem Sinne zu einer Bereicherung des Methodeninventars der Lehrevaluation bei. Gleichwohl sind aus methodologischer Sicht Einwände gegen online abgewickelte Befragungen zu erheben, die bspw. die Repräsentativität der Daten betreffen. Der Beitrag diskutiert die Chancen und Risiken, die sich durch die elektronische Evaluation -kurz: eValuation- ergeben.

1 Einleitung

Seit 1998 wird an der Universität Potsdam regelmäßig allen Lehrenden die Möglichkeit angeboten, Lehrveranstaltungen durch die Studierenden evaluieren zu lassen. Diese werden mittels eines (weitgehend standardisierten) Fragebogens befragt und gebeten, den Lehrenden ein Feedback über die von ihnen wahrgenommene Qualität der Lehrpräsentation zu geben. Dieses stellt die Grundlage für eine veranstaltungsinterne Diskussion bezüglich etwaiger Verbesserungsmöglichkeiten dar. Lehrevaluation in diesem Sinne leistet einen Beitrag zur Qualitätssicherung, indem sie als internes Optimierungsinstrument eingesetzt wird [EnKr00; EnPo01]. Voraussetzung ist, dass den Lehrenden die Ergebnisse der studentischen Befragungen möglichst zeitnah rückgemeldet werden. Nur so können die Ergebnisse noch im laufenden Semester zu einer Diskussion beitragen, bzw. als „hochschuldidaktischer Impulsgeber“ für die Planung zukünftiger Lehrveranstaltungen fungieren. Die starke Nachfrage nach den Fragebögen und der damit verbundene Aufwand, der für die Datenpflege und –auswertung zu treiben ist, hat in der Vergangenheit bewirkt, dass die Auswertungen den Lehrenden nur mit starker Verzögerung

zugegangen sind.¹ Die Evaluationsergebnisse konnten somit zu dem eigentlichen, beschriebenen Zweck nur sehr bedingt beitragen. Überdies ist durch die geschilderte Situation die Möglichkeit, den Lehrenden über deskriptivstatistische Standardauswertungen (insbesondere Mittelwerte, Häufigkeits- und prozentuale Verteilungen) hinaus, komplexeres statistisches Zahlenmaterial (bspw. Zusammenhangsmaße bezüglich verschiedener Determinanten des studentischen Qualitätsurteils) zur Verfügung zu stellen, äußerst begrenzt. Ein Ausweg aus dieser Situation ist in der Automatisierung des Procedere der Datenerhebung und ausgewählter statistischer Analysen zu sehen. Dies ermöglicht das internetgestützte Datenerhebungsverfahren eValuation, welches im vorliegenden Beitrag vorgestellt wird. Dabei wird zunächst auf einige relevante evaluationsmethodologische Grundlegungen des vorliegenden Modells der Lehrerevaluation sowie die grundsätzliche Frage nach der Qualität internetgestützter Umfrageforschung eingegangen. Beide Fragen steckten den inhaltlich-konzeptionellen Rahmen für die Realisierung des Projekts eValuation ab.

2 Evaluation der Lehre durch studentische Befragungen

Im Zuge steigender Autonomie und kompetitiver Orientierung der Hochschulen bei gleichzeitig von ihnen geforderter Rechenschaftslegung bezüglich der Verwendung öffentlicher Gelder, gewann in den 1990er Jahren auch in Deutschland die Idee einer Qualitätsüberprüfung ihrer Leistungen in Forschung sowie Studium und Lehre an Einfluss [z.B. Wiss96].² Als Instrument zur Qualitätsbeschreibung und –bewertung hat sich dabei das Forschungsprogramm der Evaluation durchgesetzt [Krom00]. Diese ist inzwischen in den Bundesländern gesetzliche Forderung [HRK01], wobei die Durchführungshoheit zumeist bei den einzelnen Universitäten liegt. Gängiges Vorgehen ist im Bereich der Lehrerevaluation neben der Erhebung von relevanten Strukturindikatoren (Ausstattungsmerkmale, Studienabbruchquoten, Berufserfolg der Absolventen und Absolventinnen, etc.) die Befragung von Studierenden. Diese werden hinsichtlich der von ihnen wahrgenommenen Qualität von Lehre und Studium insgesamt sowie der Qualität der konkreten Lehrpräsentation in den einzelnen Lehrveranstaltungen um ihre Urteile gebeten. Zur Erhebung der studentischen Lehrveranstaltungskritiken werden zumeist die Befragungstechniken der empirischen Sozialforschung eingesetzt (Interviews mittels standardisierter oder teilstandardisierter Fragebögen) [Krom00].

Der Einsatz studentischer Befragungen in der Lehrerevaluation war von Beginn an Gegenstand heftiger Kontroversen. Die methodologischen Einwände beziehen sich auf Bedrohungen der Validität studentischer Befragungsdaten [Krom95; Krom01]. Studierende seien keine kompetenten Urteiler, die Einschätzungen zumindest tendenziell von subjektiven Urteilkategorien (Bias-Variablen) verzerrt [Spie01; Wol+01]. Darüber hinaus wird die Vergleichbarkeit der studentischen Befragungen angezweifelt, die ja in jeweils verschiedenen studentischen Gruppen und unter den variierenden spezifischen

¹ Seit der Einführung des Potsdamer Modells der Lehrerevaluation durch den Akademischen Senat der Universität ist die Nutzung des Angebots der Lehrerevaluation stetig gestiegen. Derzeit werden in jedem Semester ca. 5.000 studentische Befragungen realisiert und von der zentralen „Servicestelle für Lehrerevaluation“ bearbeitet.

² Ein Überblick findet sich bei der Hochschulrektorenkonferenz [HRK00; HRK01a].

Bedingungen einer bestimmten Lehrveranstaltung erhoben werden, mithin für Leistungsvergleiche verschiedener Lehrender gänzlich ungeeignet seien. Als Gegenposition dazu wurde geäußert, dass die Studierenden als Adressaten der Lehre durchaus als relevante „Experten“ für ihre Beurteilung anzusehen und die Verzerrungen durch sachfremde Urteilskomponenten als bestenfalls moderat einzuschätzen seien [Mars84; Rind96]. Beide Positionen können sich auf einen umfangreichen Literaturkörper sowie zahlreiche empirische Studien zu diesem Thema stützen, die in den letzten Jahrzehnten entstanden sind.

Allen Bedenken zum Trotz hat sich die Methode der studentischen Lehrveranstaltungs-kritik als wichtiger Bestandteil der Lehrevaluation inzwischen in den meisten deutschen Hochschulen etabliert. Im Potsdamer Modell der Lehrevaluation wurden die genannten Einwände hinsichtlich der fehlenden Vergleichbarkeit der Daten, auf die sich etwaige Leistungsbeurteilungen von Lehrenden zu deren Nachteil stützen könnten, berücksichtigt. Die Daten werden ausschließlich den Lehrenden selbst zurückgemeldet, mit dem Auftrag (bzw. Angebot), sich mit den Studierenden bezüglich etwaiger Defizite ins Benehmen zu setzen und eine entsprechende Diskussion zu initiieren [EnPo01]. Von den Studierenden berichtete Defizite werden veranstaltungsintern behoben, die Qualität der einzelnen Lehrveranstaltungen somit auf der Durchführungsebene der Lehre gesichert. Darüber hinaus trägt diese Diskussion zur Selbstreflexion der Lehrenden und entsprechend zur etwaigen Optimierung der Planung zukünftiger Lehrveranstaltungen bei. Leistungsvergleiche zwischen verschiedenen Lehrenden (bspw. im Sinne eines Rankings) werden hingegen aus den oben genannten Gründen nicht angestrebt. Das Evaluationsmodell verfolgt einen eher sozial-kommunikativen Steuerungsansatz der Qualitätsentwicklung, indem es die Verantwortung für die Qualitätsverbesserung in die Hände der an der Lehre beteiligten Lehrenden und Lernenden legt. Dementsprechend ist die Teilnahme sowohl für die Studierenden als auch für die Lehrenden freiwillig und anonym.

Ein besonders relevantes Qualitätsmerkmal der Evaluation in diesem Modell ist seinerseits in der zeitlichen Nähe der Datenerhebung zur Auswertung und Berichterstattung zu sehen [Fuch02, S. 17]. Eine Diskussion der Evaluationsergebnisse zwischen Lehrenden und Studierenden setzt voraus, dass diese den Beteiligten noch während des laufenden Semesters zur Besprechung vorliegen. Anderenfalls bleibt die Evaluation –vermutlich zumindest in der Wahrnehmung der Studierenden– folgen- und ergebnislos. Dies wiederum könnte sich schon bald negativ in ihrer Bereitschaft niederschlagen, an der Evaluation teilzunehmen. Eine schnelle Abwicklung der Evaluationsprozedur ist daher eine Vorbedingung für ihren Erfolg. Die mit der Durchführung der Lehrevaluation beauftragte zentrale Servicestelle verfügt jedoch nur über begrenzte personelle und sächliche Ausstattungsmittel, so dass das genannte Qualitätsmerkmal der Evaluation unter den Bedingungen einer paper&pencil-basierten Befragungstechnik vielfach nicht eingehalten werden konnte.

Dies ist erst durch die Automatisierung des Verfahrens mittels der internetgestützten Befragung durch das Umfragetool eValuation, welches im Wintersemester 2003/04

erstmalig eingesetzt wurde, gelungen.³ Zusätzlich spielen jedoch auch weitere inhaltliche Überlegungen eine Rolle bei der Entscheidung, die Lehrerevaluation internetgestützt abzuwickeln. Zunächst sind bei paper&pencil Befragungen Effekte auf die soziale Erwünschtheit des studentischen Antwortverhaltens durch die Anwesenheit des Lehrenden nicht auszuschließen. Darüber hinaus, werden bei der einmalig erfolgenden Ausgabe der Fragebögen während einer Sitzung der Lehrveranstaltung die Meinungen derjenigen Studierenden nicht berücksichtigt, die diesen Termin nicht wahrnehmen können bzw. auf Grund bereits gemachter schlechter Erfahrungen nicht wahrnehmen wollen (inwieweit dies verzerrende Effekte hat, lässt sich kaum klären). Ihre Kritik ist jedoch gerade relevant für die Selbsteinschätzung und Selbstkritik der Lehrenden, die durch die studentischen Urteile unterstützt werden soll. Weiterhin besteht die Möglichkeit, dass Lehrende Fragebögen, die kritische Urteile enthalten, vor der Weitergabe an die zentral auswertende Stelle aussortieren.⁴ Schließlich ist die Zeit, die in der Lehrveranstaltung für das Ausfüllen der Fragebögen aufgewendet werden muss und damit für die Lehre nicht mehr zur Verfügung steht, als Nachteil der paper&pencil Befragungen in der Lehrerevaluation anzuführen [Fuch02].

Bevor das für die internetgestützte Datenerhebung konstruierte Umfragetool eValuation näher beschrieben wird, soll zunächst auf einige grundsätzliche Merkmale und methodologische Besonderheiten der online-basierten Umfrageforschung eingegangen werden. Diese wurden bei der inhaltlichen Konzeption von eValuation, insbesondere hinsichtlich der Einschätzung der Reichweite und Aussagefähigkeit der Daten berücksichtigt.

3 Sozialforschung online: Besonderheiten, Chancen und Risiken

EDV spielt in der empirischen Sozialforschung schon seit längerem eine zentrale Rolle bei der Verarbeitung und Auswertung statistischen Datenmaterials. Seit Ende der 1990er Jahre stellt sich für die Sozialwissenschaften zunehmend die Frage, ob Computer und Datennetze nicht darüber hinaus geeignet sind, das Methodeninventar bereits für die *Erhebung* relevanter Informationen zu erweitern. Technisch ist es inzwischen möglich, Befragungen aller Art über das Internet abzuwickeln und dabei Zeit und Kosten zu sparen. So können bspw. problemlos weltweite, kulturvergleichende Befragungen durchgeführt werden, ohne dass für diese enorme Portokosten in Kauf genommen werden müssten.

Paper&pencil Befragungen sind zudem durch das Problem des systematisierten Medienbruchs zwischen Papier- und Datenform gekennzeichnet. In vielen Fällen werden Fragebögen am Computer entworfen und anschließend durch Drucken und Kopieren vervielfältigt. Ist die Befragung abgeschlossen, so erfolgt ein erneuter Medienbruch, der

³ Nach der ersten Vorstellung des Systems im Wintersemester 2003/04 und einem Probelauf wurden auf Grund der ersten Erfahrungen entscheidende Verbesserungen vorgenommen. Diese bestehen bspw. in der Möglichkeit, die Lehrveranstaltungen und Evaluationen eigenständig anzumelden und zu administrieren.

⁴ Im beschriebenen eher sozial-kommunikativ orientierten Modell würde dies zwar hauptsächlich als eine Art Selbstbetrug zu Buche schlagen, weil die Evaluation gerade auf das Aufspüren und Beseitigen von Problemen und Defiziten ausgerichtet ist und nicht auf eine Leistungsbeurteilung der Lehrenden. Nichtsdestotrotz wäre eine entsprechende Reaktion der Lehrenden zumindest denkbar.

durch das Digitalisieren der Ergebnisse zur rechnergestützten Datenauswertung zu Stande kommt. Sofern die Ergebnisse auch in gedruckter Form veröffentlicht werden sollen, erfolgt schließlich eine dritte Umwandlung. Alle Versuche, diese Medienbrüche so „komfortabel“ und sicher wie möglich zu gestalten und somit etwaige Vorteile einer paper&pencil Befragung zu bewahren, enden in enormen finanziellen Aufwendungen für Hard- und Software, die für sehr viele Anwendungsfälle nicht aufgebracht werden können oder sich nicht rentieren. Die Vermeidung von unnötigen Medienbrüchen und das Setzen auf elektronische Lösungen und das Internet scheint zunächst dem eigentlichen Begriff der Multimedialität zu widersprechen. Allerdings bietet das Medium Internet selbst bereits die Möglichkeit, verschiedene Medientypen zu verbinden. Wo sonst werden Audio, Video und vorbereitendes Print so gut vermittelt und transportiert wie mit Hilfe des Internet?

Die besonderen Vorteile online durchgeführter Befragungen im Internet gegenüber konventionellen postalischen Befragungen bestehen im einzelnen in der *Asynchronität und Allokazität* (die Fragebögen können von den Probanden zu einer beliebigen Zeit und an einem beliebigen Ort ausgefüllt werden), *Multimedialität* (durch multimediale Elemente, wie Echtzeitstatistiken und Fortschrittsanzeigen kann Interaktivität hergestellt werden), *Objektivität der Daten* (Verzerrungen durch Interviewereffekte, wie z.B. soziale Erwünschtheit des Antwortverhaltens sowie Eingabefehler entfallen weitgehend als mögliche Fehlerquellen) und *Ökonomie* (Interviewerkosten entfallen ebenso wie Kosten für den Versand und die Retournierung der Fragebögen) [Reip03, S. 22f.; Bosn03, S. 111].

Methodologische Probleme, die im Zusammenhang mit Online-Befragungen gesehen werden, beziehen sich hauptsächlich auf die fehlende Repräsentativität der erhobenen Daten, Fragen der Stichprobenziehung sowie der fehlenden Möglichkeit der Kontrolle von Antwortausfällen (Non-Response). Repräsentativ ist eine Stichprobe dann, wenn sie aus Elementen der Grundgesamtheit besteht, für die alle die gleiche statistische Chance bestanden hat, in die Stichprobe aufgenommen zu werden [z.B. KüKr01, S. 157]. In der Praxis der empirischen Sozialforschung wird diese Bedingung durch das Ziehen von Zufallsstichproben erfüllt [z.B. Krom98, S. 258ff.]. Bezogen auf die Population der Internetnutzer ist eine Kontrolle der Zufälligkeit, mit der die Auswahl getroffen wurde (und ein verkleinertes Abbild der Grundgesamtheit abbildet) nicht möglich, weil Informationen über entscheidende Parameter der Grundgesamtheit fehlen (Anzahl aller Internetnutzer, Zusammensetzung nach Alter und Geschlecht, etc.). Es ist daher generell nicht möglich das Vorliegen einer Zufallsstichprobe anzunehmen [Pric03, S. 45]. Hinter einem an das Internet angeschlossenen Computer können sich bspw. mehrere Nutzer verbergen, verschiedene Bevölkerungsgruppen sind unterschiedlich stark im Internet repräsentiert, usw. Probleme der Stichprobenziehung beziehen sich auf die Selbstselektivität der Stichprobengenerierung [Band99]. Internetnutzer müssen die Befragung zunächst auffinden und sich darüber hinaus aktiv für eine Teilnahme entscheiden. Inwiefern die Prozesse, die zu dieser Entscheidung führen, verzerrend wirken (stark positive oder negative Einstellung gegenüber dem Befragungsgegenstand) kann vom Untersuchungsleiter kaum kontrolliert werden. Die Schwierigkeit von Antwortausfällen (Non-Response) besteht insbesondere in der Verkleinerung der Stichprobe [Schn97]. In der Regel soll von den in der Stichprobe gemessenen Parametern, unter Inkaufnahme einer gewissen statistischen Unsicherheit auf die entsprechenden unbekannt Parameter der

Grundgesamtheit geschlossen werden, aus der die Stichprobe gezogen wurde [z.B. Krom98, S. 248]. Das Ausmaß der Streuung der Stichprobenparameter um den entsprechenden Parameter der Grundgesamtheit wird dabei von zwei Größen bestimmt: der Varianz in der Grundgesamtheit und der Stichprobengröße [Schn97]. Im Fall von online durchgeführten Befragungen spielt die Verkleinerung der Stichprobe zwar wegen der generell vergleichsweise großen Datensätze vermutlich nur eine untergeordnete Rolle. Problematisch ist vielmehr, dass nicht ausgeschlossen werden kann, dass die Ausfälle systematisch erfolgen. Der Beweis, dass Ausfälle nicht systematisch, sondern zufällig erfolgen, ist nicht zu erbringen, denn dazu wären ja gerade Informationen nötig, die ausfallbedingt fehlen. Systematische Ausfälle können jedoch zu erheblichen Verzerrungen führen, etwa, wenn der Erhebungsgegenstand Mobilität und der Ausfallgrund Immobilität ist [Schn97].

Da für Internetbefragungen weder geprüft werden kann, ob die Auswahl eine Zufallsstichprobe ist, noch ob Ausfälle (die ja ihrerseits bspw. durch Nichtkenntnis der Befragung entstehen) zufällig sind, muss von Verzerrungen der erhobenen Daten zumindest ausgegangen werden.⁵ Sofern das Ziel der geplanten Erhebung darin besteht, Aussagen über eine Grundgesamtheit durch die Merkmale einer untersuchten Stichprobe zu begründen, sollte der Einsatz internetgestützter Befragungstechniken daher eher mit Skepsis betrachtet werden [Haupt99].⁶

In weniger grundlagenforscherisch bzw. „bevölkerungsrepräsentativ“ orientierten Anwendungsfällen empirischer Forschung [Heck03, S. 83], in denen die genannten methodologischen Fragen der Repräsentativität eine eher untergeordnete Rolle spielen, muss auf die entscheidenden Vorteile der neuen Medien dagegen nicht verzichtet werden. Insbesondere trifft dies zu, wenn die Grundgesamtheit der möglichen Fälle überschaubar groß ist und die zu befragenden Personen kontrolliert ausgewählt und – bspw. per email- angesprochen und um das Ausfüllen eines online vorgehaltenen Fragebogens gebeten werden können. Bei derartigen Einsatzfeldern von Onlineforschung kann es sich bspw. um Mitarbeiterbefragungen in Unternehmen handeln [Heck03, S. 87] oder eben auch um die Befragung von Studierenden hinsichtlich ihrer Qualitätseinschätzungen zu besuchten Lehrveranstaltungen. Studierende bilden eine der Bevölkerungsgruppen, die das Internet am intensivsten nutzen. Mehr als 90 % der Studierenden verfügen über einen Internetanschluss und ein E-Mailkonto [Band03, S. 80; BMBF02]. An der Universität Potsdam bspw. wird die kostenlose Einrichtung eines persönlichen Accounts für alle Studierenden angeboten. Die Nutzung der neuen Medien für die Lehrveranstaltungsverwaltung (bspw. als elektronische Veranstaltungsbelegung mittels des „Blackboard Learning System“, bei der auch Emailverteilerlisten der teilnehmenden Studierenden angelegt werden können) findet immer weitere Verbreitung im Bewusstsein der Lehrenden und Studierenden. Dementsprechend kann sich die Teilnehmerrekrutierung für die internetgestützte Lehrerevaluation auf eine nahezu lückenlos „versorgte“ und informierte Nutzerschaft stützen. Diese günstigen

⁵ Das Problem von (Non)Response ist derzeit ein zentrales Thema der Methodenentwicklung innerhalb der Onlineforschung [z.B. Bosn02; Bosn03, S. 116ff.]

⁶ Um die Vorteile von Internetbefragungen (komplikationsloses Erreichen großer Datensatzzahlen) doch für die empirische Sozialforschung nutzbar zu machen, werden verschiedene Wege gegangen. So werden bspw. online durchgeführte Bevölkerungsumfragen durch kleinere thematisch identische „konventionelle“ (z.B. postalische) Befragungen flankiert, um entsprechende Gewichtungen vornehmen zu können [Pers04].

Diese günstigen Ausgangsbedingungen wurden für die Konstruktion und Konzeption des Umfragetools eValuation genutzt.

4 eValuation: Umfragetool für die Lehrevaluation

Die erhobenen Daten, die sich auf die Beurteilung einzelner Lehrveranstaltungen beziehen, sollen nicht für einen Repräsentativitätsschluss auf die Qualitätseinschätzungen aller Potsdamer Studierenden (bzw. der Studierenden einer Fakultät oder eines Faches) herangezogen werden, sondern der veranstaltungsinternen Qualitätssicherung dienen. Dementsprechend ist dieser Aspekt der Datenqualität im vorliegenden Anwendungsfall nicht virulent. Dem Vorwurf der mangelnden Repräsentativität der Daten durch die zu erwartende (und auch tatsächlich eingetretene) sinkende Teilnahmebereitschaft der Studierenden (s.u.) kann daher entgegengehalten werden, dass die Daten *insbesondere* die Qualitätsurteile derjenigen Studierenden widerspiegeln sollen, die eine substanzielle Kritik (positiver oder negativer Färbung) vorzubringen haben und durch ihre Teilnahme zum Ausdruck bringen, dass sie an der Qualitätssicherung der Lehre interessiert mitwirken wollen. Als Argument wäre anzuführen, dass eValuation insbesondere solche Studierenden zur Teilnahme anregt, die sich aktiv mit dem Geschehen der Lehrveranstaltung auseinandersetzen und das entsprechende Interesse durch die Bereitschaft dokumentieren, auch außerhalb der Lehrveranstaltung „Online-Zeit“ zu investieren. Von dieser Gruppe Studierender könnten dann auch substanzielle Kritiken und Beurteilungen erwartet werden, die den Prozess der lehrevaluation-internen Qualitätssicherung und -entwicklung nachhaltig unterstützen. Die Frage der Repräsentativität spielt dann eine nicht mehr so herausragende Rolle für den Erfolg der Lehrevaluation, wenn die Lehrenden auf Grund der vorgetragenen Kritik zur Selbstreflexion angeregt werden und die studentischen Urteile bspw. durch den Besuch hochschuldidaktischer Fortbildungen umzusetzen versuchen.

eValuation versucht diesen Prozess zu unterstützen, indem es den Lehrenden die Möglichkeit bietet, selbständig Lehrveranstaltungen zur Bewertung durch die Studierenden anzumelden und zu verwalten. Der Zeitraum, der für die Bewertung von den Lehrenden der Veranstaltungen vorgesehen wird, kann durch die Lehrenden selbst bestimmt werden. Das System gibt lediglich Rahmentermine vor, die sich an Anfangs- und Schlusstermin der Vorlesungszeit orientieren. Nach Abschluss der Evaluation werden automatisch anfordernde statistische Berechnungen durchgeführt und die Auswertungsergebnisse, unterstützt durch Tabellen und Diagramme, für Einsicht und Druck online zur Verfügung gestellt. Der Zugriff auf die Ergebnisse steht nur dem jeweiligen Lehrenden offen und ist dementsprechend passwortgeschützt. Den Lehrenden steht jedoch die Option frei, den teilnehmenden Studierenden ebenfalls online Zugriff auf die Auswertungsergebnisse zu gewähren und somit einen entscheidenden Beitrag zum Nutzen und zur Akzeptanz der studentischen Veranstaltungskritik zu erbringen.

Mit einer Kennziffer greifen die Studierenden zuvor auf den online vorgehaltenen Fragebogen zu und geben die Veranstaltungsbewertungen ab. Gegenüber dem herkömmlichen Verfahren hat sich inhaltlich nichts geändert, es wird der gleiche standardisierte Fragebogen eingesetzt und zusätzlich die Gelegenheit gegeben, offen weitere Kritiken

und Anregungen zu formulieren. Diese gehen dem Lehrenden unmittelbar per Email zu und werden von der zentralen Servicestelle nicht ausgewertet. Sie dienen ausschließlich der Information der Lehrenden.

Die augenfälligen Vorteile von eValuation für die Lehrenden und Studierenden gegenüber dem herkömmlichen paper&pencil Modell bestehen in

- dem zeitnahen Vorliegen der Auswertungsergebnisse und der entsprechenden Verfügbarkeit für die Diskussion mit den Studierenden,
- der Vereinfachung des Procedere durch das Entfallen der Notwendigkeit, Fragebögen zu bestellen, zu verteilen und an die auswertende zentrale Servicestelle zurückschicken,
- der größeren Anonymität und einer „entspannteren“ Bewertungssituation während des Ausfüllens der Fragebögen,
- der Möglichkeit für Lehrende, differenziertere statistische Auswertungen anzufordern, welche spezifische individuelle Informationsbedürfnisse befriedigen und auf Wunsch von der zentralen Servicestelle durchgeführt werden können.

Durch den Einsatz der neuen Medien in der Lehrerevaluation kann diese – eben durch das zeitnahe Vorliegen der Auswertungsergebnisse - in höherem Maße ihrem Anspruch gerecht werden, als internes Optimierungsinstrument die problembezogene Diskussion Lehrender und Studierender zu initiieren. Auf diese Weise trägt eValuation dazu bei, dass die Lehrerevaluation nicht bloße „Datenfriedhöfe“ erzeugt, sondern tatsächliche Qualitätsverbesserungen mittels der beschriebenen sozialkommunikativen Herangehensweise anschiebt.

5 Technische Umsetzung von eValuation

Bei eValuation handelt es sich um eine typische Client/Server Anwendung mit einer webbasierten Benutzerschnittstelle. Das heißt, dass die Interaktion zwischen Anwender und Software mit Hilfe eines Browsers per Internet erfolgt. Diese Musterarchitektur wurde gewählt, weil sie für die Anwender von eValuation, Lehrende und Studierende, die flexibelste Lösung darstellt und die technischen Voraussetzungen zur Realisierung durch die Universität Potsdam gegeben sind. eValuation wurde unter Verwendung der Skriptsprache PHP und dem Datenbankmanagementsystem MySQL realisiert. Damit ist sowohl die clientseitige Anwendbarkeit per Browser als auch die serverseitige Lauffähigkeit auf allen in diesem Bereich gängigen Betriebssystemen und Basisinstallationen gesichert.

Die Realisierung der internetgestützten Abwicklung der studentischen Lehrveranstaltungs-kritik an der Universität Potsdam kann in vier Bereiche gegliedert werden: (1) die Verwaltung von Lehrenden und deren Lehrveranstaltungen sowie von relevanten Daten Studierender, (2) die Realisierung der internetgestützten Datenerhebung mit Hilfe eines

elektronischen Fragebogens, (3) die Auswertung der erhobenen Daten und (4) die Aufbereitung und Darstellung der Ergebnisse. Auf die Besonderheiten der genannten Anwendungsbereiche soll im Folgenden näher eingegangen werden.

- (1) Die Benutzer von eValuation sind zum einen Dozentinnen und Dozenten (Lehrende) und zum anderen Studentinnen und Studenten (Studierende). Um einen geordneten und gesicherten Ablauf der Anwendung zu gewährleisten, müssen Lehrende zur Benutzung von eValuation autorisiert sein und vom System authentifiziert werden können. So lange die Universität keine semesteraktuellen Lehrendendaten in digitaler Form bereitstellen kann, erfolgt die Prüfung der sich anmeldenden Personen manuell durch die Servicestelle für Lehrevaluation. Dies führt momentan zu einem minimalen Verzug zwischen Registrierung und Freischaltung. Durch die Mitteilung des zur Nutzung notwendigen Passwortes über eine gültige Emailadresse des Lehrenden wird möglichem Missbrauch vorgebeugt.

Lehrende verwalten ihre im Zuge der Evaluation zu beurteilenden Lehrveranstaltungen mit Hilfe der Anwendung selbst. Dabei haben sie die Möglichkeit, Profile für die einzelnen Veranstaltungen wie üblich zu erstellen (Eingabe von Titel, Veranstaltungsart, Ort und Zeit der Lehrveranstaltung), zu bearbeiten und zu löschen. Mit Hilfe des Evaluationsstatus einer Veranstaltung steuern die Lehrenden u.a. den Zeitraum der Bewertung. Veranstaltungen können zur Bewertung durch die Studierenden offen stehen (Status: aktiv), die Bewertung kann für das laufende Semester abgeschlossen sein (Status: beendet) oder Veranstaltungen können von der Evaluation im laufenden Semester ausgeschlossen sein (Status: inaktiv). Unmittelbar nach Beenden des Bewertungszeitraumes (Status: beendet) stehen den Lehrenden – und, sofern von ihnen vorgesehen, auch den Studierenden – die Ergebnisse zur Verfügung. Somit wird für Lehrende eine höchstmögliche Flexibilität bei der Gestaltung der Veranstaltungskritik und sowohl für Lehrende als auch für Studierende eine zeitnahe Information über die Ergebnisse ermöglicht.

Die im Vordergrund stehenden funktionellen Anforderungen an eValuation machen eine Verwaltung von Studierendendaten nicht erforderlich. Auch aufgrund des Datenschutzes muss die Anonymität der Studierenden gewährleistet bleiben. Dennoch spielt die Möglichkeit eines anonymisierten Wiedererkennens von Studierenden, bspw. im Fall von Wiederholungsbefragungen, eine auch aus methodischer oder forschersicher Hinsicht wichtige Rolle.⁷

Weil das Ausfüllen des Fragebogens mit Hilfe von eValuation nicht mehr in den Lehrveranstaltungen selbst sondern zu einer beliebigen anderen Zeit online geschieht, müssen die Studierenden vom System als Besucher bestimmter Lehrveranstaltungen authentifiziert werden können. Dies erfolgt mit der Eingabe der vom Lehrenden für die jeweils besuchte Lehrveranstaltung mitgeteilten Kennnummer. Die Kommunikation zwischen Lehrenden und Studierenden über

⁷ Bspw. im Falle von etwaigen Wiederholungsbefragungen ist über die Möglichkeit einer anonymisierten Wiedererkennung nachzudenken.

den Zeitraum der Bewertung und über das Erreichen des elektronischen Fragebogens wird somit zu einer der wichtigsten Voraussetzungen. An vielen Lehrstühlen der Universität Potsdam erfolgt die Einschreibung in Lehrveranstaltungen bereits elektronisch. Zur lehrveranstaltungsbezogenen Kommunikation werden in diesem Rahmen zumeist auch die Emailadressen der Studierenden erhoben (z.B. zur Mitteilung über den Ausfall von Veranstaltungen aufgrund von Krankheit des Lehrenden). Die erhobenen Emailadressen können auch zur Aufforderung der Studierenden, sich an der Evaluation zu beteiligen genutzt werden.

- (2) Der bei der Bewertung von Lehrveranstaltungen zu Grunde liegende Fragebogen ist inhaltlich identisch mit dem Fragebogen, der bei der auch weiterhin parallel angebotenen Erhebung des studentischen Feedbacks per paper&pencil Methode verwendet wird. Lediglich die Verwendung einer Software zur Erstellung und anschließend automatisierten Auswertung des klassisch auszufüllenden Fragebogens erfordert einige Designunterschiede. Prinzipiell können die klassisch und elektronisch erhobenen Daten miteinander verglichen werden (vgl. Abschnitt 6.1).

Der elektronische Fragebogen zeichnet sich gegenüber seinem klassischen Pendant jedoch durch eine höhere Benutzerfreundlichkeit aus. So steht beim Aufruf bereits fest, zu welchem Lehrenden die Veranstaltung gehört und welchen Typs diese ist. Auf diese Art und Weise können bspw. Fragen, die nur für einen bestimmten Veranstaltungstyp relevant sind, automatisch aus- und eingeblendet werden.

Das Ausfüllen des Fragebogens mit Hilfe des Internets wirft das Problem der Kontrolle mehrfachen Abschickens der einmal eingegebenen Daten auf. Würde dies uneingeschränkt möglich gemacht werden, so käme es vermutlich häufig zu einer Verzerrung zumindest der häufigkeitsbezogenen Ergebnisse. Auch wenn es derzeit aus technologischer Sicht keine zu 100 Prozent sichere Methode gibt, dieses Problem mit einem für diesen Rahmen vertretbaren Aufwand zu beherrschen, ohne dass die Gefahr bestünde, dass Studierende von der Bewertungsmöglichkeit ausgeschlossen würden (z.B. durch IP/Zeit-Kontrollen), so trifft eValuation doch Vorkehrungen, um allzu leichte Manipulationen zu verhindern.

- (3) Die im Rahmen von eValuation standardmäßige Auswertung der erhobenen Daten (Deskriptivstatistik, Häufigkeiten, Mittelwerte u.a.) erfolgt automatisiert. Weiterführende komplexere Auswertungen können nach individueller Absprache mit der Servicestelle für Lehrerevaluation realisiert werden. Die zur automatisierten Auswertung erforderlichen statistischen Berechnungen werden durch die Möglichkeiten des DBMS MySQL unterstützt.

- (4) Die Präsentation der Bewertungsergebnisse erfolgt auf den Internetseiten von eValuation. Für Lehrende und Studierende⁸ stehen jedoch auch speziell für den Druck aufbereitete Versionen bereit. Die Visualisierung erfolgt in Form von Tabellen und Diagrammen. Die Erzeugung letzterer wird durch die Verwendung einer für diese Zwecke dienenden, frei verfügbaren Klassenbibliothek unterstützt. Entsprechende serverseitige Voraussetzungen konnten nach Absprache mit der Universität Potsdam geschaffen werden.

6 Erste Erfahrungen

Im Wintersemester 2003/04 wurde eValuation erstmals und parallel zu der herkömmlichen Befragungsmethode (paper&pencil) angeboten und von unerwartet vielen Lehrenden genutzt. Insgesamt meldeten sich gut 140 Dozenten und Dozentinnen bei eValuation an, ca. 190 Lehrveranstaltungen wurden online evaluiert und damit insgesamt mehr als 1.500 studentische online abgewickelte Befragungen realisiert.⁹ Erwartungsgemäß sank die Beteiligungsquote der Studierenden zunächst teilweise beträchtlich ab.¹⁰ Im Fall der herkömmlichen Befragungsmethode wird die Lehrveranstaltung für einige Zeit unterbrochen, in der den Studierenden das Ausfüllen der Fragebögen ermöglicht wird. Die überwiegende Mehrheit nimmt in dieser Konstellation an der Befragung teil, die Frage der Rücklaufquote stellt sich nicht, weil durchgängig nahezu Vollerhebungen realisiert werden können [Gau+04]. Im Falle der online erhobenen Urteile wird von den Studierenden verlangt, selbständig auf den im Internet vorgehaltenen Fragebogen zuzugreifen. Dies setzt wiederum voraus, dass Online-Zeit investiert wird in das Feedback, welches zunächst möglicherweise als hauptsächlich dem Lehrenden nützlich wahrgenommen wird. Dass die Beteiligungsquote erhöht werden kann, wenn die Ansprache der Studierenden erfolgt, wenn sie ohnehin mit der Lehrveranstaltung beschäftigt sind (bspw. durch das Downloaden von Unterrichtsmaterialien aus dem Internet) oder sie die Bitte um Mitwirkung per Email über die lehrveranstaltungseigene Verteilerliste erhalten, lässt sich zwar vermuten und aus anderen Untersuchungen zeigen [z.B. Gau+04], eine gesicherte diesbezügliche Empirie auf Grund der an Universität Potsdam bislang erhobenen Daten liegt jedoch noch nicht vor.

6.1 Vergleich online vs. offline Erhebungen

Die parallele Erhebung der studentischen Qualitätsurteile mit zwei Befragungsmethoden im Semester der Einführung von eValuation (online und offline, bzw. paper&pencil-

⁸ Sofern die Lehrenden dies vorsehen werden den Studierenden die Auswertungen in einer Druckversion als Download zur Verfügung gestellt.

⁹ Dem stehen ca. 3.500 Befragungen gegenüber, die mittels der konventionellen paper&pencil Befragung durchgeführt wurden. Inwieweit dieses Übergewicht auf die Skepsis der Lehrenden gegenüber eValuation oder andere Faktoren (z.B. Unkenntnis des neuen Angebots) zurückzuführen ist lässt sich bislang nicht quantifizieren.

¹⁰ In einigen ausgewählten Lehrveranstaltungen wurde sowohl die Zahl der insgesamt teilnehmenden Studierenden als auch die Zahl der sich an der Lehrveranstaltung beteiligenden Studierenden ermittelt. Die Beteiligungsquote variierte hier zwischen 27 und 60 Prozent.

basiert) macht einen ersten Vergleich der Ergebnisse möglich. Allerdings muss dieser mit der gebotenen Vorsicht zur Kenntnis genommen werden. Es können nämlich nur jeweils unterschiedliche Lehrveranstaltungen von verschiedenen Lehrenden und verschiedenen Beurteilergruppen miteinander verglichen werden. Ein Vergleich von ähnlichen Seminaren eines Lehrenden (z.B.: Übungsgruppe 1 eines Seminars evaluiert online, Übungsgruppe 2 offline) oder von zwei unterschiedlichen Veranstaltungen durch die gleiche Studierendengruppe mit jeweils einer der Erhebungsmethoden ließ sich nicht realisieren. Es können daher nur alle online durchgeführten mit allen offline durchgeführten Befragungen verglichen werden.

Um etwaige Unterschiede zwischen den Ergebnissen der Befragungsarten festzustellen, wurden Mittelwertvergleiche durchgeführt.¹¹ Da die erhobenen Daten in den einzelnen Items signifikant von der Normalverteilung abweichen, wurden nichtparametrische U-Tests durchgeführt.¹² Wider Erwarten ließen sich zwischen den Befragungsmethoden Unterschiede in den studentischen Urteilen feststellen. Da sich beide Gruppen Befragter aus der Gruppe der Studierenden der Universität Potsdam rekrutierten und mithin hinsichtlich demographischer Merkmale (Geschlecht, Alter, Bildungsstatus) keine relevanten Unterschiede erwartet werden dürften, wurde davon ausgegangen, dass die unterschiedlichen Erhebungsmethoden gleiche oder zumindest nicht übermäßig variierende Ergebnisse produzieren würden [vgl. Band+01; Fuch02]. Die Ergebnisse zeigen jedoch, dass die online befragten Studierenden hinsichtlich des wahrgenommenen Erkenntnis- und Motivationszuwachses, der durch den Besuch der jeweiligen Lehrveranstaltung entstanden ist, zu einem positiveren Ergebnis kommen als die offline Befragten ($p = .000$). Items, die das „Lernklima“ der Veranstaltung zum Gegenstand haben, werden ebenfalls unterschiedlich beurteilt. So werden die Transparenz der Lernanforderungen, die Ansprechbarkeit der Lehrenden, die Beratung in Fragen des Leistungsscheinwerbs sowie die Einschätzung des studentischen Engagements für das Gelingen der Lehrveranstaltung positiver eingeschätzt ($p \leq .016$). Das Tempo des Lehrvortrages und die Frage, ob sich ein roter Faden durch den Ablauf der Lehrveranstaltung ziehe wird dagegen von den online Befragten kritischer beurteilt als von der Referenzgruppe ($p = .000$).

Relativiert werden diese festgestellten Unterschiede, wenn man den Vergleich nach einzelnen Fächern getrennt vornimmt. So lösen sich im Fall der Studierenden der Wirtschaftswissenschaften (die im Untersuchungszeitraum die anteilig größte Gruppe der an der Evaluation teilnehmenden Studierenden stellten) die Unterschiede zwischen den Einschätzungen der online und offline befragten Personen größtenteils auf. Die Einschätzung hinsichtlich Erkenntnis- und Motivationsgewinn sowie verschiedene Fragen zum Lernklima weisen keine signifikanten Unterschiede zwischen den Erhebungsmodi auf. Dies legt den Schluss nahe, dass beide Verfahren die studentischen Urteile in gleicher Weise abbilden können [Fuch02, S.12]. Die festgestellten Unterschiede lassen sich möglicherweise auch auf generell unterschiedliche Urteilmuster zurückführen, die über

¹¹ Diese beziehen sich auf die vorgegebenen 7stufigen Rating-Skalen zur Beurteilung verschiedener Aspekte der wahrgenommenen Qualität der Lehrpräsentation („Fragen zum Lernklima“).

¹² Die Items bilden das studentische Urteil auf Skalen ab, die bspw. zwischen den Polen „sehr gut“ und „sehr schlecht“ oder „stimme gar nicht zu“ und „stimme voll und ganz zu“ rangieren. Da das studentische Urteil generell eher positiv ausfällt, liegen durchgängig schiefe Verteilungen vor (Kolgomorov-Smirnoff-Test; $p \leq .001$), die die Durchführung eines üblicherweise für Mittelwertvergleiche gewählten t-Tests ausschließen.

die Fächer variieren. Wie bereits erwähnt, schließen die Vergleiche jeweils alle (in allen Fächern) online bzw. offline erhobenen Daten ein. Eine genauere Klärung dieser Frage wird Gegenstand weiterer Forschungen sein, soll jedoch im Rahmen des vorliegenden Beitrags nicht weiter verfolgt werden.

7 Ausblick und weitere Entwicklungsmöglichkeiten

Die ersten Erfahrungen – breite Beteiligung der Lehrenden, Verkürzung der benötigten Zeit für die Evaluationsprozedur – legen den Schluss nahe, dass mit der Entwicklung von eEvaluation eine Verbesserung der Lehrevaluation erreicht werden konnte. Die Frage der sinkenden Teilnahmebereitschaft der Studierenden wird sich vermutlich in den kommenden Jahren als nicht mehr virulent erweisen, wenn man eine Entwicklung hin zur flächendeckenden Ausbreitung eines generell elektronischen Lehrveranstaltungsmanagements unterstellt, welches die unkomplizierte Ansprache der Studierenden via E-mail-Verteilerliste ermöglicht. Weitere Entwicklungsmöglichkeiten des Tools sind jedoch in der Integration neuer zusätzlicher Funktionen zu sehen. So ist die Schaffung eines „Online-Forums eEvaluation“ geplant, welches im Sinne einer Metaevaluation den Nutzern die Möglichkeit geben soll, Wünsche und Anregungen zur Durchführung der Lehrevaluation zu diskutieren. Auch ist über ein teilweises Individualisieren der Fragebögen und die Schaffung der Option für die Lehrenden, eigene Erhebungsinstrumente, gemäß (lehrveranstaltungs-)spezifischer Informationsbedürfnisse, zusammenzustellen, nachzudenken. Zwar sollte die Möglichkeit des Vergleichs der Evaluationsergebnisse einzelner Lehrveranstaltungen nicht gänzlich aus der Hand gegeben werden, vordringlichstes Ziel der Lehrevaluation im hier beschriebenen Evaluationsmodell ist jedoch weniger der (Leistungs-)Vergleich unter den Lehrenden, als ihre selbstkritische Auseinandersetzung mit den eigenen Leistungen in Lehre und Studium.

Literaturverzeichnis

- [Band99] Bandilla, W.: WWW-Umfragen – Eine alternative Datenerhebungstechnik für die empirische Sozialforschung? In: Batinic, B.; Werner, A.; Gräf, L.; Bandilla, W. (Hrsg.): Online Research. Methoden, Anwendungen und Ergebnisse. S. 9-20. Göttingen u.a.: Hogrefe. 1999.
- [Band03] Bandilla, W.: Die Internet-Gemeinde als Grundgesamtheit. In: Informationszentrum Sozialwissenschaften (Hrsg.): Online Erhebungen. Reihe *Sozialwissenschaftliche Tagungsberichte*, Band 7. S. 71-81. Bonn. 2003.
- [Ban+01] Bandilla, W.; Bosnjak, M.; Altdorfer, P.: Effekte des Erhebungsverfahrens? Ein Vergleich zwischen einer web-basierten und einer schriftlichen Befragung zum ISSP-Modul Umwelt. ZUMA-Nachrichten. Jg. 25 (49). S. 7-28. Mannheim. 2001.
- [BMBF02] Bundesministerium für Bildung und Forschung (BMBF): Computernutzung und neue Medien im Studium. Ergebnisse der 16. Sozialerhebung des Deutschen Studentenwerks. Bonn. 2002.
- [Bosn02] Bosnjak, M.: (Non)Response bei Web-Befragungen- Auswahl, Erweiterung und empirische Überprüfung eines handlungstheoretischen Modells zur Vorhersage und Erklärung des Partizipationsverhaltens bei web-basierten Fragebogenuntersuchungen. *Berichte aus der Psychologie*. Aachen: Shaker. 2002

- [Bosn03] Bosnjak, M.: Web-basierte Fragebogenuntersuchungen- Methodische Möglichkeiten, aktuelle Themen und Erweiterungen. In: Informationszentrum Sozialwissenschaften (Hrsg.): Online Erhebungen. Reihe *Sozialwissenschaftliche Tagungsberichte*, Band 7. S. 109-133. Bonn. 2003.
- [EnKr00] Engel, U.; Krekeler, G.: Der Wechsel vom Grund- ins Hauptstudium. Studentische Einschätzungen zu Beginn des Hauptstudiums an der Universität Potsdam. Reihe Ergebnisberichte zur Evaluation, Bd. 3. Herausgegeben v. d. Universität Potsdam/Arbeitsgruppe „Qualität der Lehre“. Potsdam. 2000.
- [EnPo01] Engel, U.; Pohlenz, P.: Lehre und Studium im Spiegel studentischer Bewertungen. Über das Potsdamer Modell der Lehrerevaluation. In: Spiel, C. (Hrsg.): Evaluation universitärer Lehre –zwischen Qualitätsmanagement und Selbstzweck. S. 131-150. Münster: Waxmann. 2001
- [Fuch02] Fuchs, M.: Internet-basierte Lehr-Evaluation. Erste Ergebnisse und Einladung zur Mitwirkung. *Soziologie*. Heft 2, S. 5-18. 2002.
- [Gau+04] Gau, W. ; Bomhardt, C. ; Schmidt-Mänz, N.: Einsatz von computergestützter Lehrveranstaltungsevaluation. Zeitschrift für Evaluation. 1/04. S. 35-50. Wiesbaden. 2004.
- [Haupt99] Hauptmanns, P.: Grenzen und Chancen von quantitativen Befragungen mit Hilfe des Internet. In: Batinic, B.; Werner, A.; Gräf, L.; Bandilla, W. (Hrsg.): Online Research. Methoden, Anwendungen und Ergebnisse. S. 21-38. Göttingen u.a.: Hogrefe. 1999.
- [Heck03] Heckel, C.: Online gewonnene Stichproben – Möglichkeiten und Grenzen. In: Informationszentrum Sozialwissenschaften (Hrsg.): Online Erhebungen. Reihe *Sozialwissenschaftliche Tagungsberichte*, Band 7. S. 83-94. Bonn. 2003.
- [HRK00] HRK (Hochschulrektorenkonferenz): Wegweiser 2000 durch die Qualitätssicherung von Lehre und Studium. *Dokumente und Informationen* (2). Bonn. 2001.
- [HRK01] HRK: Hochschulgesetzliche Regelungen zur Qualitätssicherung. Hochschulrahmengesetz und Hochschulgesetze der Länder (Stand 31.05.2001). *Beiträge zur Hochschulpolitik* (7). Bonn. 2001.
- [HRK01a] HRK: Zukunftsaufgabe Qualitätsentwicklung. Drei Jahre „Projekt Q“. Beiträge zur Hochschulpolitik (4). Bonn. 2001.
- [Krom95] Kromrey, H.: Evaluation durch Umfrageforschung? Methodische Fallstricke bei der Messung von Lehrqualität durch Befragung von Vorlesungsteilnehmern. In: Mohler, P. (Hrsg.): Universität und Lehre. Ihre Evaluation als Herausforderung an die empirische Sozialforschung. S. 105-128. Münster: Waxmann. 1995
- [Krom98] Kromrey, H.: Empirische Sozialforschung. 8. Aufl. Opladen: Leske und Budrich. 1998.
- [Krom00] Kromrey, H.: Qualität und Evaluation im System Hochschule. In: Stockmann, R. (Hrsg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. S. 234-258. Opladen: Leske u. Budrich. 2000.
- [Krom01] Kromrey, H.: Studierendenbefragungen als Evaluation der Lehre? Anforderungen an Methodik und Design. In: Engel, U. (Hrsg.): Hochschulranking. Zur Qualitätsbewertung von Studium und Lehre. S. 11-48. Frankfurt/Main: Campus. 2001.
- [KüKr01] Kühnel, S.-M.; Krebs, D.: Statistik für die Sozialwissenschaften. Grundlagen, Methoden, Anwendungen. Reinbek: Rowohlt. 2001.
- [Mars84] Marsh, H. W.: Students' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases and Utility. *Journal of Educational Psychology*, Vol. 76 (5), 707-754. 1984.
- [Pers04] Perspektive Deutschland: Methodik. Wie repräsentativ sind die Ergebnisse von Perspektive Deutschland? www.perspektive-deutschland.de/02050.php (zuletzt: August 2004).
- [Pric03] Pricking, T.: Online-Erhebungen in der amtlichen Statistik: Das Internet als neues Medium für die Erhebung und Übermittlung statistischer Daten. In: Informationszentrum Sozialwissenschaften (Hrsg.): Online Erhebungen. Reihe *Sozialwissenschaftliche Tagungsberichte*, Band 7. S. 43-60. Bonn. 2003.

- [Reip03] Reips, U.-D.: Online-Erhebungen in der wissenschaftlichen Sozialforschung. In: Informationszentrum Sozialwissenschaften (Hrsg.): Online Erhebungen. Reihe *Sozialwissenschaftliche Tagungsberichte*, Band 7. S. 21-30. Bonn. 2003.
- [Rind96] Rindermann, H.: Zur Qualität studentischer Lehrveranstaltungsevaluation. Eine Antwort auf die Kritik an der Lehrevaluation. *Zeitschrift für Pädagogische Psychologie*, 10 (3-4). S. 129-145. 1996.
- [Schn97] Schnell, R.: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske und Budrich. 1997.
- [Spie01] Spiel, C.: Der differentielle Einfluss von Biasvariablen auf studentische Lehrveranstaltungsbeurteilungen. In: Engel, U. (Hrsg.): Hochschulranking. Zur Qualitätsbewertung von Studium und Lehre. S. 61-82. Frankfurt/Main: Campus. 2001.
- [Wiss96] Wissenschaftsrat: Empfehlung zur Stärkung der Lehre in den Universitäten durch Evaluation. Bonn. 1996.
- [Wo+01] Wolf, P.; Spiel, C.; Pellert, A.: Entwicklung eines Fragebogens zur globalen Lehrveranstaltungsevaluation – ein Balanceakt zwischen theoretischem Anspruch, Praktikabilität und Akzeptanz. In: Spiel, C. (Hrsg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck. S. 89-110. Münster: Waxmann. 2001.

