aus: Zeitschrift für Evaluation, 1/2004, S. 35 – 50

# Einsatz von computergestützter Lehrveranstaltungsevaluation

*Wolfgang Gaul, Christian Bomhardt, Nadine Schmidt-Mänz*
*Institut für Entscheidungstheorie und Unternehmensforschung,*
*Universität Karlsruhe (TH)*

## 1. Einleitung

Seit Anfang der 90er Jahre wird insbesondere Qualitätssicherung als wichtige Aufgabe für die weitere Entwicklung der Hochschulen empfohlen (siehe DUI: 7). Dies führte zur Gründung verschiedener Evaluationsinstitutionen (z.B. 1994: Verbund Norddeutscher Universitäten (siehe VND), 1995: Zentrale Evaluationsagentur (siehe ZEvA), 1997: Evaluationsagenturen Nordrhein-Westfalen (siehe Rindermann 2001b)). Der Gesetzgeber verankerte Evaluationen durch das vierte Gesetz zur Änderung des Hochschulrahmengesetzes (siehe HRG) vom 24.02.1997. Im neuen § 6 heißt es: „Die Arbeit der Hochschulen in Forschung und Lehre soll regelmäßig bewertet werden. Die Studierenden sind bei der Bewertung der Qualität der Lehre zu beteiligen. Die Ergebnisse der Bewertung sollen veröffentlicht werden". Diesen Punkten wird durch Lehrveranstaltungsbeurteilungen im Rahmen von Evaluationsverfahren Rechnung getragen (siehe EvaNet).

Aus übergeordneter Sicht ist hierzu Folgendes anzumerken: Wenn in einer Gesellschaft Angebote für erkannte Bedürfnisse (hier: Durchführung von computergestützten Lehrevaluationen) neu auf den Markt kommen, kann man oft beobachten, dass sich erst ein so genanntes „dominantes Design" im Rahmen der Angebotsvielfalt durchsetzen muss. Gerade für computergestützte Angebote ist das Ausmaß, bis zu dem Einigkeit über Standardisierungsmöglichkeiten herrscht, ein wichtiger Gesichtspunkt. Für den vorliegenden Beitrag ist zu betonen, dass die Unterstützung von Lehrevaluationen mit Hilfe eines computergestützten Systems (mit den Design-Nebenbedingungen: „Notebook-geeignet" und „Internet-fähig") im Vordergrund stand. Um die Akzeptanz des Systems durch potentielle Nutzer zu erhöhen, wurden auf der Grundlage einschlägiger Evaluationsliteratur verschiedene Module mit zugehörigen Evaluationsdimensionen zur Auswahl bereitgestellt, aus denen sich in Zusammenarbeit mit so genannten Leadusern des Systems auf Basis von aus der Evaluationsforschung bekannten Fragebögen akzeptierte Module herauskristallisierten, die zusätzlich Kriterien wie Vergleichbarkeit und Nachhaltigkeit zu erfüllen hatten. Vergleichbarkeit und Nachhaltigkeit sind weitere Gründe für eine Standardisierung der Fragebögen und Nutzung von Datenbank-Funktionalitäten, wie sie im Rahmen computergestützter Evaluation bereitgestellt wird.

Im nächsten Abschnitt 2 werden zunächst wesentliche Eigenschaften von papiergebundener und computergestützter Evaluation von Lehrveranstaltungen angesprochen, bevor in Abschnitt 3 auf eine Auswahl von Software, die für computergestützte Lehrveranstaltungsevaluationen Unterstützung anbieten kann, eingegangen wird, um zu verdeutlichen, welche Anforderungen an computergestützte Lehrevaluationssysteme zu berücksichtigen sind. Abschnitt 4 ist der Beschreibung des eigenen Systems ETU NetEval (Institut für *E*ntscheidungs*T*heorie und *U*nternehmensforschung, Inter*NetEval*uation), welches sich z.Z. an der Universität Karlsruhe in der Erprobungsphase befindet, und der Darstellung erster Ergebnisse aufgrund der bisher durchgeführten Probeläufe gewidmet. In Abschnitt 5 erfolgt ein Ausblick auf zukünftige Entwicklungen.

## 2. Papiergebundene vs. computergestützte Evaluation von Lehrveranstaltungen

Ziele und Rahmenbedingungen von Lehrveranstaltungsevaluationen beeinflussen deren Ausgestaltung (siehe Rindermann 2001a) unabhängig davon, für welche Evaluationsform – papiergebunden oder computergestützt – man sich entscheidet. Exemplarisch sei hier auf die Aktivitäten zweier Fakultäten an der Universität Karlsruhe hingewiesen.

Die Fakultät für Elektrotechnik lässt sich seit einiger Zeit durch das amerikanische Accreditation Board for Engineering and Technology (siehe ABET) evaluieren. Dabei wurde ein mehrstufiger Qualitätskontrollprozess (Befragung von Studierenden, Absolventen, Doktoranden, Ehemaligen und Industrievertretern) zur Verbesserung von Forschung und Lehre eingeführt. Für die Lehrveranstaltungsevaluation wird ein standardisierter Fragebogen verwandt, der – geeignet ausgefüllt – maschinengelesen werden kann und auf dieser Basis Standardauswertungen erlaubt.

Bei der bisher papiergebundenen Lehrevaluation an der Fakultät für Wirtschaftswissenschaften wurden die Studierenden durch die Fachschaft mit mehrseitigen Fragebögen zu ausgewählten Lehrveranstaltungen einmal während des Semesters befragt. Die Fragebögen wurden manuell in einer Datenbank erfasst, ausgewertet und veröffentlicht. Die Veröffentlichung erfolgte zu einem Zeitpunkt, an dem die evaluierte Veranstaltung bereits vorüber war. Die Ergebnisse der Evaluation konnten so nur für spätere Veranstaltungen in qualitätsentwickelndem Sinne genutzt werden.

Beispiele wie die soeben geschilderten lassen erkennen, welche Probleme in Evaluationssituationen auftreten können. Schwierigkeiten bei der Beschaffung und Auswertung von Fragebögen bei papiergebundenen Evaluationen sowie evtl. bei der Veröffentlichung der Ergebnisse können abgemildert werden, weil teilweise zentrale Stellen an den verschiedenen Hochschulen, wie z.B. die Geschäftsstelle Evaluation der Universität Leipzig, vorhanden sind, die abhängig von den Evaluationszielen und Rahmenbedingungen u.a. die Fragebogenbeschaffung durch standardisierte Fragebögen, ergänzbare Kernfragebögen oder Fragen- sowie Layoutvorlagen unterstützen und bei der Auswertung ausgewählter Informationen helfen. In der Regel bleibt es Evaluatoren allerdings unbenommen, ihre Fragebögen selbst zusammenzustellen und sie in Eigenregie auszuwerten, wenn man sich dadurch größere Flexi-

bilität und bessere Erreichbarkeit der Evaluationsziele verspricht. Um zusätzliche Fragen ergänzte Kernfragebögen bilden bei der papiergebundenen Evaluation einen Kompromiss zwischen standardisierten und veranstaltungsspezifischen Fragebogenvarianten. Allerdings müssen die Fragebögen der Geschäftsstelle Evaluation der Universität Leipzig beispielsweise zentral beschafft werden, da selbst Ausdrucke von Fragebogenvorlagen im PDF-Format für eine maschinelle Auswertung nicht exakt genug sind (siehe ULGE).

Zu den Vorteilen einer papiergebundenen Lehrveranstaltungsevaluation gehört die Möglichkeit, die Fragebögen während der Veranstaltung parallel ausfüllen zu lassen, sodass diese nur kurz unterbrochen werden muss. Ein Nachteil kann sein, dass eine zeitnahe Auswertung der Evaluationsergebnisse nicht möglich ist. Falls Fragebögen standardisiert von zentraler Stelle vorgehalten werden, um eine schnelle Auswertbarkeit zu unterstützen, können veranstaltungsspezifische Evaluationsfragen nicht berücksichtigt werden. Ergänzungen eines Kernfragebogens oder die Entwicklung eines komplett neuen Fragebogens sind mit entsprechendem Aufwand verbunden. Der arbeitsintensivste Teil einer papiergebundenen Evaluation (siehe Schomann et al. 2000) entsteht oft durch die Übertragung der erhobenen Umfragedaten in eine geeignete Datenbank, von der aus eine Auswertung der Ergebnisse erfolgen kann und evtl. Ausgestaltungswünsche im Hinblick auf die graphische Aufbereitung wichtiger Resultate berücksichtigt werden können. Zwischen papiergebundener Evaluation und Auswertung kann so ein längerer Zeitraum vergehen, so dass eine Besprechung der Ergebnisse oder eine mögliche Integration des Feedbacks aus der Evaluation im Rahmen der laufenden Veranstaltung nicht mehr möglich ist.

Ein nahe liegender Gedanke ist der Einsatz von Software, mit der man u.a. auch die angesprochenen störenden Limitierungen papiergebundener Lehrevaluationen aufheben will. Da im Rahmen computergestützter Lehrevaluation unterschiedliche, ursprünglich auch für andere Zwecke konzipierte Computerprogramme in Betracht kommen können, wollen wir unter computergestützten Lehrevaluationssystemen solche Software verstehen, die jeweils einen Bereich für registrierte *Evaluatoren* zum Anlegen und Auswerten von Evaluationsfragebögen und einen Bereich für (registrierbare) Teilnehmer an der Lehrveranstaltungsevaluation, kurz: *Umfrageteilnehmer*, zum Beantworten der Evaluationsfragen umfaßt. Gängige Systeme enthalten ferner einen Bereich für *Administratoren*. Um eine (sinnvolle) ortsunabhängige Nutzung dieser Bereiche zu ermöglichen, müssen zeitgemäße Systeme Internet-Technologien unterstützen. Computergestützte Systeme können allen drei zuvor genannten Nutzergruppen verschiedene Vorteile bieten. Evaluatoren können die Unterlagen für die Evaluation einfacher und schneller erstellen und sofort auswerten. Umfrageteilnehmer können während der Veranstaltung mit Notebooks über Funknetzwerk oder unabhängig von der Veranstaltung (z.B. in Rechnerpools) per Computer an der Evaluation teilnehmen und zeitnah (evtl. beim nächsten Lehrveranstaltungstermin) ausgewählte Evaluationsergebnisse kennenlernen und mit den Evaluatoren diskutieren. Dabei können auch Fragen berücksichtigt werden, mit denen sich der individuelle Stand der Leistung und Wissensaufnahme von Studierenden mit aggregierten Werten unterschiedlicher Veranstaltungsteilnehmergruppierungen (z.B. Spitzengruppe vs. Hinteres Drittel) vergleichen lässt bzw. es kann überprüft werden, wie stark die eigene Meinung mit den für unterschiedliche Gruppierungen aggregierten Werten übereinstimmt. Schließlich können Administratoren Umfrageteilnehmer und Eva-

luatoren flexibler unterstützen und gleichzeitig übergeordnete Evaluationsziele wie Vergleichbarkeit und Nachhaltigkeit besser verfolgen.

### Fazit:

Papiergebundene Evaluation: Insgesamt erzeugen standardisierte Fragebögen bei der papiergebundenen Lehrveranstaltungsevaluation den geringsten Aufwand. Mit programmierten Scannern können die Fragebogeninhalte (evtl. nur teilweise) in Datenbanken übertragen werden. Da standardisierte Fragebögen jedoch nur mit Aufwand umgestaltet und an individuelle Vorstellungen angepasst werden können (Probleme z.B. bei Layoutänderungen, beim Neudruck, bei der Maschinenlesbarkeit), ist zwischen Änderungsaufwand und Erzeugung/Erhöhung von Akzeptanz abzuwägen (vgl. Souvignier/Gold 2003). Kernfragebögen können durch zeitintensive Layoutarbeiten um veranstaltungsspezifische Teile ergänzt werden, allerdings können heterogenes Fragebogenlayout, unterschiedliche Formulierungen oder die Verwendung verschiedener Skalen für inhaltlich ähnliche Fragen aufgrund des Einsatzes von Vorlagen aus mehreren geeignet erscheinenden Evaluationsmaterialien die Bereitschaft zur Umfrageteilnahme senken und bei der Auswertung zu Fehlern führen. Auswertungen von papiergebundenen Evaluationen sind erst nach einer von der Erfassungsmethode, Ressourcenauslastung und Anzahl der Antworten abhängigen Wartezeit verfügbar.

Computergestützte Evaluation: Fragebogengestalterische und organisatorische Einschränkungen wie bei papiergebundenen Evaluationen entfallen weitgehend, nachdem Layoutaspekte geklärt sind. Kernfragebögen können – sofern vom System ermöglicht – schnell um eigene Fragen ergänzt werden, da das Layout automatisch vom System angepasst wird. Die Erstellung individueller Fragebögen kann durch eine integrierte Fragensammlung oder durch Fragenblöcke bezüglich verschiedener Evaluationsdimensionen erleichtert werden. Zusammen mit einem vorgegebenen, unter ergonomischen Aspekten entwickelten Layout (strukturierte Anordnung, homogene Farb- und Symbolwahl, wenig Skalenwechsel, alle Antwortmöglichkeiten sind direkt ersichtlich) und unter Berücksichtigung von Akzeptanzaspekten können optisch ansprechende und visuell schnell erfassbare Fragebögen erzeugt werden. Der gesamte Evaluationsablauf kann durch den Einsatz von Computersystemen beschleunigt und vereinfacht werden. Die sofortige Verfügbarkeit von Auswertungen und der mögliche Export der Umfragerohdaten für individuelle Auswertungen autorisierter Nutzer sind weitere Vorteile.

## 3. Auswahl von Software zur Gestaltung von Lehrveranstaltungsevaluationen

### 3.1 Kriterien zur Beschreibung computergestützter Lehrevaluationssysteme

Computergestützte Hilfsmittel zur Lehrevaluation können durch Kriterien wie z.B. (in verkürzter Form ausgedrückt) *Anpassbarkeit*, *(Umfrageteilnehmer-)Autorisierung*, *Vorlagenpool*, *Auswertung* und *Ergonomie* kategorisiert werden.

Anpassbarkeit umfasst Unterdimensionen wie *Layout*, die Gestaltbarkeit von *Antwortskalen* und Flexibilität bei der Einbeziehung von *Fragen*. Die verschiedenen Ausgestaltungsmöglichkeiten von Software für Lehrveranstaltungsevaluationen bezüglich Anpassbarkeit resultieren in diversen Arten der Gestaltbarkeit von Fragebögen durch die Evaluatoren (i.d.R. sind das die Dozenten). Allerdings bekommt bei Softwaresystemen das Kriterium Autorisierung ein besonderes Gewicht. Evaluatoren und Administratoren werden dabei üblicherweise durch eine Kombination aus Benutzername und Passwort autorisiert, die u.a. die Rechte zur Weiterverarbeitung der bei Evaluationen anfallenden Daten regelt. Umfrageteilnehmer können durch eine Umfrage-Nummer und ein dazu passendes Passwort autorisiert werden. Umfrageteilnehmerautorisierung kann durch die Ausprägungen „Keine Autorisierung" (es ist kein Passwort notwendig, jeder kann beliebig oft an der Evaluationsumfrage teilnehmen), „Globale Autorisierung" (ein Passwort ermöglicht die mehrfache Teilnahme an der Evaluation, dadurch können einzelne Personen, die das Passwort kennen, mehrfach Feedback abgeben) und „Individuelle Autorisierung" (jeder Umfrageteilnehmer erhält ein individuelles Passwort, welches das Ausfüllen von genau einem Evaluationsfragebogen ermöglicht) beschrieben werden. Falls vorhanden, kann ein Vorlagenpool Evaluatoren durch eine Sammlung von Beispielfragebögen, vorgefertigte Fragemodule, einzelne Evaluationsfragen und die Bereitstellung von Antwortskalen sowie Beispiellayouts im Rahmen der beim Kriterium Anpassbarkeit diskutierten Unterdimensionen unterstützen. Das Kriterium Auswertung beschreibt die von der entsprechenden Software angebotenen Möglichkeiten zur Analyse der erhobenen Umfragedaten. Schließlich soll das Kriterium Ergonomie die Eignung von Software für Lehrveranstaltungsevaluationen unter dem Gesichtspunkt der einfachen und intuitiven Bedienbarkeit charakterisieren helfen.

## 3.2 Auswahl von Software zur Lehrveranstaltungsevaluation

Grundsätzlich können Lehrveranstaltungsevaluationen an mehreren Stellen durch Software vereinfacht werden. Dies kann z.B. durch Programme zum Planen von Evaluationen, zum Gestalten von Evaluationsfragebögen, zum Auswerten von Evaluationsdaten etc. geschehen. Vor dem Hintergrund der soeben diskutierten Kriterien haben wir uns auf Software konzentriert, welche eine Benutzerschnittstelle für Evaluatoren (zum Anlegen und Auswerten von Lehrveranstaltungsevaluationen) und für Umfrageteilnehmer (zum Ausfüllen von Evaluationsfragebögen) enthält, per Internet-Technolgien ortsunabhängig genutzt werden kann und typischerweise auch eine Benutzerschnittstelle für Administratoren besitzt und damit eine Evaluation vollständig und nicht nur stellenweise unterstützt. Um einen Eindruck von angebotenen computerbasierten Hilfen für Lehrveranstaltungsevaluationen zu erhalten, wurden vier Softwarepakete ausgewählt, darunter zwei Systeme, die ausschließlich für computergestützte Evaluationen verwendet werden (Inquiry der FH Weingarten (siehe INQUIRY) und LIQUIS der Universität Linz (siehe LIQUIS)). Desweiteren enthalten e-Learning-Systeme teilweise ein Evaluationsmodul: Wir haben exemplarisch das CLIX-System (siehe CLIX) untersucht. Software für webbasierte Umfragen kann ebenfalls zu Evaluationszwecken verwendet werden, wofür wir exemplarisch FormGen (siehe FORMGEN) ausgewählt haben.

*Tabelle 1*: Auswahl von Software zur Unterstützung computerbasierter Lehrevaluationen

| Software: | | Inquiry | LIQUIS | CLIX | FormGen |
|---|---|---|---|---|---|
| Anpassbarkeit | Layout | vorgegeben | vorgegeben | vorgegeben | individuell gestaltbar |
| | Antwortskalen | vorgegeben | vorgegeben | individuell gestaltbar | individuell gestaltbar |
| | Fragen | individuell gestaltbar | individuelle Zusatzfragen | individuell gestaltbar | individuell gestaltbar |
| Umfrageteilnehmerautorisierung | | keine Autorisierung oder individuelle Autorisierung | individuelle Autorisierung | individuelle Autorisierung durch e-Learning System | keine Autorisierung oder individuelle Autorisierung |
| Vorlagenpool | | nicht vorhanden | Kernfragebögen | nicht vorhanden | nur für Layouts |
| Auswertung | | Häufigkeits-verteilungen mit Mittelwert, Anzeige der Umfagedaten in Tabellenform | Häufigkeits-verteilungen mit Mittelwert, Anzeige der einzelnen Fragebögen | Häufigkeits-verteilungen mit Mittelwert, Varianz und Standard-abweichung | Häufigkeits-verteilungen, Datenexport |
| Ergonomie | | verbesserungs-würdig | gut | attraktiv aber ermüdend | gut |

Die Auswahl der untersuchten Software soll andeuten, welches Spektrum an computerbasierten Unterstützungsmöglichkeiten, die u.U. aufgrund anderer Hauptzielsetzungen entwickelt wurden, für die Ausgestaltung computergestützter Lehrevaluationssoftware von Interesse sein kann. Tabelle 1 zeigt die Ausgestaltung der zuvor diskutierten Kriterien für die ausgewählte Software zur computerbasierten Lehrveranstaltungsevaluation. Das Layout kann ausschließlich bei FormGen individuell gestaltet werden. Die Antwortskalen der Fragen können bei CLIX und FormGen angepasst werden. Inquiry, CLIX und FormGen erlauben individuelle Fragen. LIQUIS bietet verschiedene Kernfragebögen an, von denen einer ausgewählt werden muss. Dieser kann um individuelle Fragen ergänzt werden. Alle vier untersuchten Softwarepakete bieten individuelle Umfrageteilnehmerautorisierung an – FormGen und Inquiry ermöglichen wahlweise Umfragen ohne Umfrageteilnehmerautorisierung, für LIQUIS und CLIX ist die individuelle Autorisierung obligatorisch. Die individuelle Autorisierung wird bei Inquiry durch das Versenden einer e-Mail, die einen Link auf die Umfrage und ein individuelles Passwort enthält, realisiert. FormGen und LIQUIS realisieren die individuelle Autorisierung durch Erzeugen einer Liste mit individuellen Passwörtern, die der Evaluator ausdrucken, ausschneiden und an die Umfrageteilnehmer weitergeben kann. Bei CLIX erfolgt die individuelle Autorisierung durch Anmeldung am e-Learning-System. Bei LIQUIS kann die Autorisierung alternativ über eine Kombination von Matrikel-Nr. und Kennwort erfolgen. An Vorlagen bietet LIQUIS Kernfragebögen an, die um einzelne Fragen ergänzt werden können. Bei den anderen Programmen sind weder Fragen- noch Umfragevorla-

gen vorhanden. Bezüglich des Kriteriums Auswertung liefern alle Systeme Häufigkeitsverteilungen, teilweise mit weiteren Kennzahlen (Mittelwert und Standardabweichung oder Varianz). Inquiry und LIQUIS stellen auf Wunsch die ausgefüllten Fragebögen einzelner Umfrageteilnehmer dar. FormGen kann die Umfragerohdaten exportieren. Die Bedienung ist bei Inquiry sowohl für Evaluatoren als auch für Umfrageteilnehmer verbesserungsbedürftig (nicht besonders übersichtlich, sehr viele Klicks bis zum fertigen Ergebnis). CLIX verwendet attraktive aber ermüdende Benutzerschnittstellen (Jede Frage wird bei der Umfrageteilnahme einzeln abgefragt, sehr viele Klicks bis zum fertigen Ergebnis.). Die Benutzerschnittstellen von FormGen und LIQUIS sind insgesamt attraktiv, intuitiv und effizient handhabbar.

## 3.3 Schwachstellen bei einer computergestützten Lehrveranstaltungsevaluation

Bei papiergestützter Evaluation war es an der Fakultät für Wirtschaftswissenschaften der Universität Karlsruhe üblich, die Vorlesung für eine angemessene Zeit zu unterbrechen und die Studierenden in diesem Zeitraum die Evaluationsfragebögen ausfüllen zu lassen, d.h. (fast?) alle anwesenden Studierenden nahmen an der Evaluation teil. Rücklaufquoten waren nicht von Interesse. Bei der computergestützten Lehrveranstaltungsevaluation kann die Durchführung dergestalt erfolgen, dass sich die Studierenden außerhalb der Lehrveranstaltung an einen Rechner setzen, um die Fragebögen auszufüllen. Dies kann zu Problemen in Bezug auf die Repräsentativität der Umfrageteilnehmer im Verhältnis zu allen Teilnehmern an der Lehrveranstaltung führen. Dass ein Umfrageteilnehmer mehrfach antwortet, kann bei computergestützten Lehrveranstaltungsevaluationen durch geeignete Autorisierung vermieden werden. Andererseits hat man sich darüber Gedanken zu machen, wie man möglichst viele Lehrveranstaltungsteilnehmer auch zur Teilnahme an der Evaluation motiviert. Gute Rücklaufquoten können erreicht werden, wenn die Aufforderung zur Teilnahme an der Evaluation den Studierenden präsentiert wird, während sie sich sowieso mit der Veranstaltung beschäftigen (Hinweise während der Lehrveranstaltung, den Übungen oder Tutorien sowie beim Download von Veranstaltungsmaterialien). Umfragen, die die Eigeninitiative der Teilnehmer zum Ausfüllen erfordern, führen zu niedrigeren Rücklaufquoten. Falls Studenten, wie es bei der zuvor diskutierten Software – mit Ausnahme von CLIX – der Fall ist, motiviert werden müssen, sich aktiv an der Veranstaltungsevaluation zu beteiligen, sinken die Rücklaufquoten. Durch eine strikte Autorisierung können sich die Rücklaufquoten weiter verschlechtern, ebenso durch eine als nicht anonym empfundene Autorisierung, die zusätzlich zu einer Verfälschung der Ergebnisse führen kann. Bzgl. der Akzeptanz von Seiten der Evaluatoren ist anzumerken, dass verschiedene fortgeschrittene Anwender gemäß unserer Erfahrungen eine größere Vielfalt und eine erweiterte Auswahl in Bezug auf Auswertungsmöglichkeiten wünschen. Versierte Anwender von Datenanalyse-Systemen (Power-User) können mit ihren bevorzugten Softwarepaketen natürlich beliebige Auswertungen durchführen, falls die Umfragerohdaten in elektronischer Form vorliegen und in gewünschter Weise portierbar sind. Diesem wichtigen Akzeptanzaspekt wird von der zuvor angesprochenen Software lediglich FormGen gerecht.

# 4. Das eigene System ETU NetEval und Einsatzerfahrungen

## 4.1 Ausgestaltung von ETU NetEval

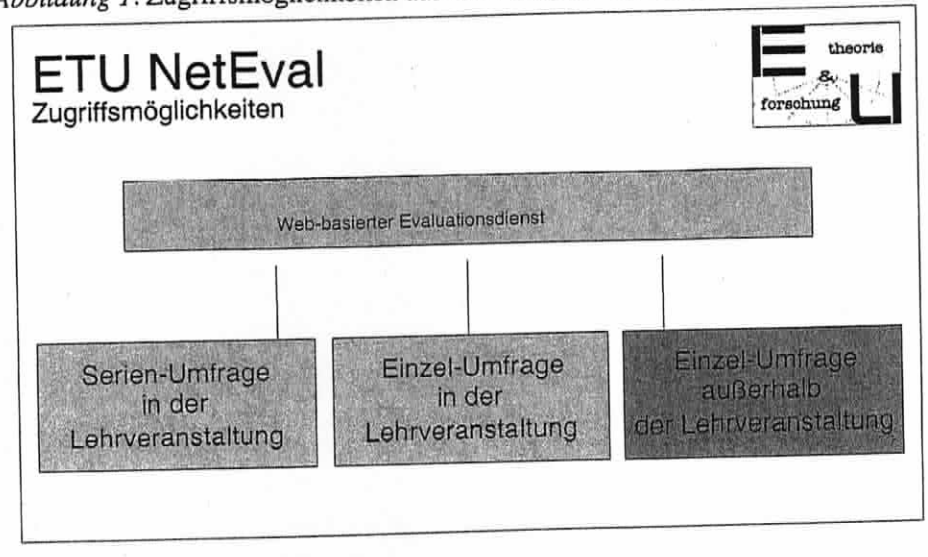
Im Rahmen eines gemeinsamen Teilprojekts zusammen mit anderen Forschern wurde für die Notebook Universität Karlsruhe (TH) (siehe NUKATH) am Aufbau von Evaluationsdiensten für mobile Lernumgebungen gearbeitet, wobei zwei Richtungen verfolgt wurden. Einerseits wurde am Rechenzentrum der Universität Karlsruhe vorhandene Anwendungssoftware (Application Hosting) in einem konzeptionellen Architekturrahmen zu einem Evaluationssystem zusammengefügt, bei dem die Schnittstellenproblematik noch nicht vollständig gelöst ist. Daran wird z.Z. gearbeitet. Andererseits entstand das ETU NetEval System in Eigenregie, weil kurzfristig erste Erfahrungen mit computergestützten Lehrveranstaltungsevaluationen gesammelt werden mussten.

Das ETU NetEval System enthält jeweils eine Benutzerschnittstelle für Evaluatoren, Umfrageteilnehmer und Administratoren. Mit der Administrationsschnittstelle kann die Rechteverwaltung durchgeführt werden, neue Dienstenutzer können freigeschaltet, erlaubte Kombinationen von Fragen bzw. Frageblöcken administriert und weitere Fragebögen als Blöcke hinzugefügt werden. ETU NetEval enthält Frageblöcke für verschiedene Evaluationsarten (steuernd, qualitätsentwickelnd) und Veranstaltungstypen. Diese wurden unter Berücksichtigung der aktuellen Evaluationsforschung eingepflegt (siehe Kromrey 2001; Mitterauer 1996; Rindermann 2001b; Schratz 1996; Souvignier/Gold 2003). Umfrageteilnehmer werden im aktuellen Prototyp über ein Umfragepasswort autorisiert, wobei zwischen globaler Autorisierung in der Veranstaltung auf Basis kontrollierter Weitergabe der eingesetzten Notebooks und individueller Autorisierung außerhalb der Veranstaltung zu unterscheiden ist. Die Schnittstelle für Evaluatoren erlaubt ein schnelles und einfaches Anlegen von Veranstaltungsevaluationen. Auswertungen können unmittelbar nach Beendigung der Evaluation abgerufen werden, Umfragerohdaten können für weiterreichende Auswertungen mit geeigneter Software (z.B. Excel, SPSS) exportiert werden.

ETU NetEval verwendet Internet-Technologien und profitiert von dem an der Universität Karlsruhe vorhandenen Funknetzwerk (siehe DUKATH). Mit dem campusweit verfügbaren Funknetzwerk können die Studierenden langfristig Lehrevaluationen auf Notebooks mit Funknetzanschluss überall auf dem Campus, d.h. auch innerhalb der jeweiligen Lehrveranstaltung, vornehmen. Bis sich Notebook-Veranstaltungen auf breiter Front durchgesetzt haben, werden die Evaluationen in „kleineren" Lehrveranstaltungen als „Serien-Umfrage in der Lehrveranstaltung" (siehe Abbildung 1) mit wenigen Notebooks, die in der Veranstaltung von Studierendem zu Studierendem weitergegeben werden, durchgeführt. Bei dieser Nutzungsart präsentiert ETU NetEval nach Beendigung eines individuellen Evaluationsvorgangs eine neue Einstiegsseite für den nächsten Nutzer, an den das Notebook weitergereicht wird. Diese Serien-Umfrage-Methode führte zu mit papiergebundener Evaluation vergleichbaren Rücklaufquoten. Alternativ kann auf lehrveranstaltungsbegleitenden Webseiten ein Link zur Umfrage gesetzt werden oder es können die Zugangsdaten zur Beteiligung an der Evaluation individuell aufbereitet zusammen mit vorlesungsbegleitenden Materialien angegeben werden, um Evaluationen auch außerhalb von Lehrveranstaltungen durchführen zu lassen.

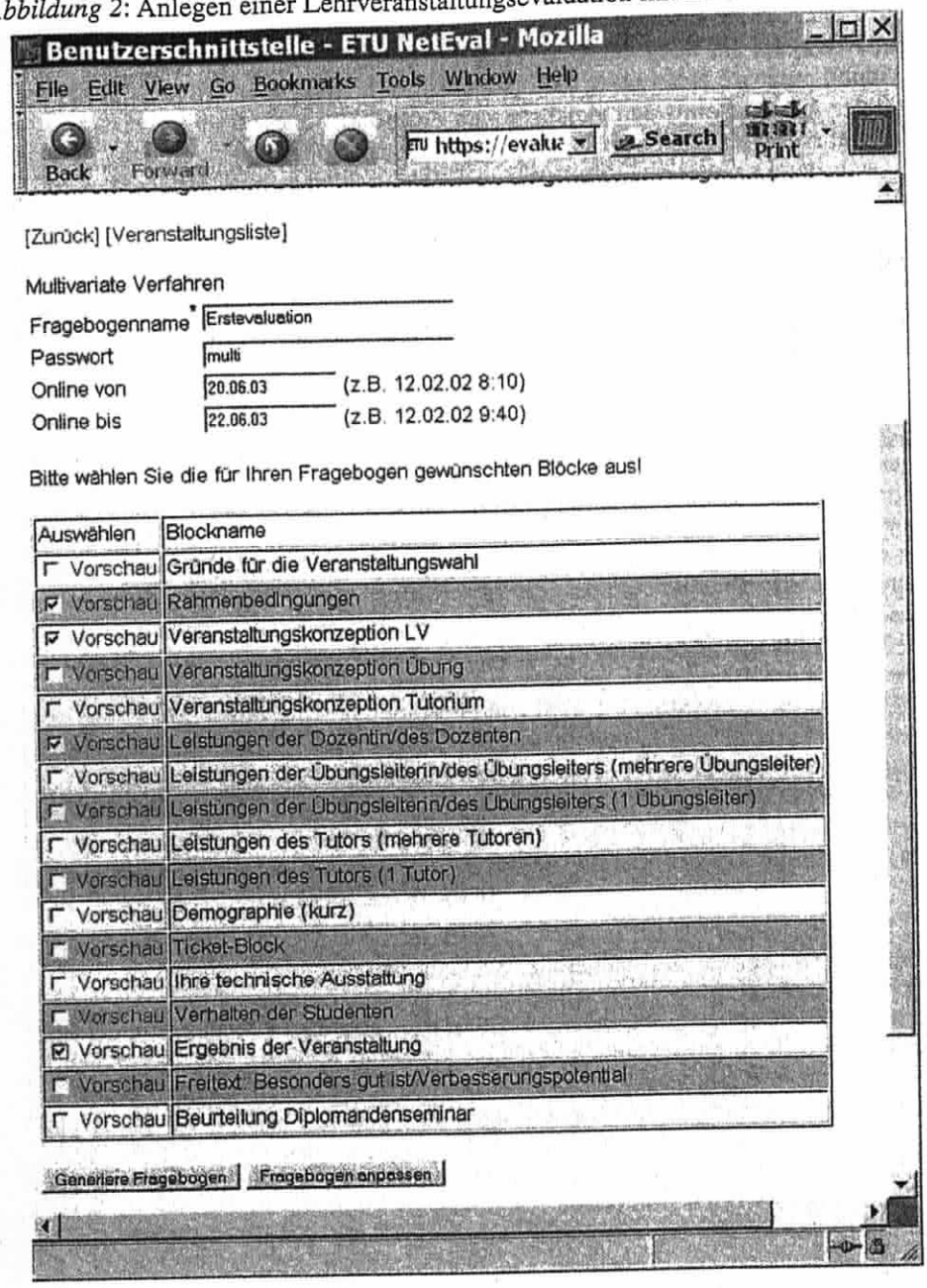
Wenn es in Zukunft auf dem Campus genug Access-Points (Übergangspunkte vom Funknetz ins Campusnetz) gibt, um aus jeder Lehrveranstaltung unmittelbar über Notebook auf das Internet zugreifen zu können, wenn alle Studenten über individuell nutzbare Notebooks mit Funknetzanschluss verfügen und wenn die Kapazität der Access-Points den gleichzeitigen Zugriff auf das Internet auch aus „größeren" Lehrveranstaltungen erlaubt, wird die Zugriffsmöglichkeit mittels „Einzelumfrage in der Lehrveranstaltung", die in ETU NetEval bereits vorgesehen ist, voll zum Einsatz kommen können. Dann werden sofort nach der Evaluation die Ergebnisse in geeignet aufbereiteter Form während der Lehrveranstaltung vorliegen und diskutiert werden können.

*Abbildung 1*: Zugriffsmöglichkeiten auf ETU NetEval im Überblick




Mit ETU NetEval sollte ein möglichst modernes System geschaffen werden, das für qualitätsentwickelnde und steuernde Evaluationen genutzt werden kann. Für ein solches System sind u.a. folgende Überlegungen zu berücksichtigen: Es muss (1) auch von nicht-computerversierten Interessenten, die eine Evaluation durchführen wollen, einfach zu nutzen sein, soll (2) ein schnelles Anlegen von geeigneten Fragen aus vorgefertigten Fragenblöcken ermöglichen (siehe Abbildung 2), soll (3) in verschiedenen Umfragen identische Fragen verwenden, muss (4) für die Evaluationsproblematik ausgewählte Fragedimensionen beantworten und soll (5) Fragebögen so darstellen, dass sie für die Umfrageteilnehmer einfach, übersichtlich und schnell zu beantworten sind (siehe Abbildung 3). Einfache Bedienbarkeit, schnelle Durchführbarkeit und ergonomische Gestaltung sollen primär für Akzeptanz sorgen. Überlegungen zur Ergonomie und Nachhaltigkeit erfordern eine systemweit einheitliche Ausgestaltung von Frageblöcken, ein standardisiertes Layout sowie feststehende, in späteren Evaluationen wiederkehrende Fragen, die die Vergleichbarkeit mit früheren Ergebnissen sicherstellen.

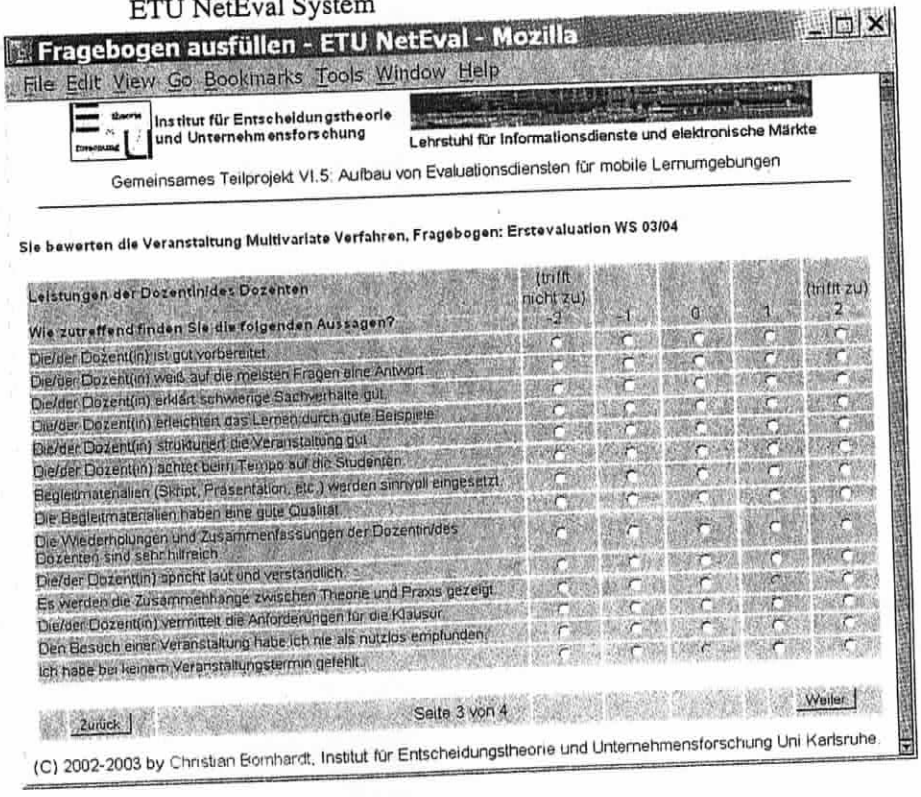
*Abbildung 2*: Anlegen einer Lehrveranstaltungsevaluation mit ETU NetEval




Die Zusammenstellung von unterschiedlich formulierten aber inhaltlich ähnlichen Fragen kann auch bei Wahlmöglichkeiten aus verschiedenen Frageblöcken durch das System verhindert werden. Zur einfachen Handhabung und zum Abdecken von relevanten Evaluationsdimensionen wurden vorgefertigte Fragenblöcke angeboten.

Abbildung 2 zeigt den Stand (Dezember 2003) der Auswahlmöglichkeiten von Fragebl\u00f6cken in ETU NetEval. Im dort wiedergegebenen Beispiel wurden die Blöcke „Rahmenbedingungen", „Veranstaltungskonzeption LV", „Leistungen der Dozentin/des Dozenten" und „Ergebnis der Veranstaltung" zur Auswahl angeklickt.

*Abbildung 3*: Beispiel des Frageblocks „Leistungen der Dozentin/des Dozenten" im ETU NetEval System




In Abbildung 3 wird beispielhaft der Fragenblock „Leistungen der Dozentin/des Dozenten" präsentiert, der im Rahmen der Darstellung erster praktischer Erfahrungen mit ETU NetEval für weitere Erläuterungen benutzt wird (siehe Abbildung 4). Für vergleichende Evaluationen können einzelne Frageblöcke aus einem komplett standardisierten Fragebogen bestehen oder Fragen zu einzelnen ausgewählten Evaluationsdimensionen enthalten. Zur Erhöhung der Flexibilität können für qualitätsentwickelnde Umfragen einzelne Fragen eines Frageblocks abgewählt werden.

## 4.2 Praktische Erfahrungen mit ETU NetEval

Mittlerweile sind eine Reihe von Lehrveranstaltungen mittels ETU NetEval evaluiert worden. Tabelle 2 gibt einen Überblick unterteilt nach „Serien-Umfrage in der Lehrveranstaltung" und „Einzel-Umfrage außerhalb der Lehrveranstaltung". Die dritte

Möglichkeit „Einzel-Umfrage in der Lehrveranstaltung" wird z.Z. im internen Umfeld geprüft. Die durchschnittliche Dauer der Beschäftigung eines Studierenden mit der Evaluation im Rahmen der Serien-Umfrage in der Lehrveranstaltung lag bei der aktuell benutzten Version bei ca. 5½ Minuten, allerdings ergab sich wegen der limitierten Anzahl der einsetzbaren Notebooks in Abhängigkeit von der Größe der Veranstaltung ein Zeitrahmen von bis zu einer Stunde, in der die Notebooks weitergegeben wurden.

*Tabelle 2*: Überblick über die bisher mit ETU NetEval durchgeführten Evaluationen (Dezember 2003)

| Vorlesung | Semester | Hörer | Teilnehmer | Quote | Anzahl | Anzahl | Φ Dauer pro Fragebogen | Umfragegesamtdauer |
|---|---|---|---|---|---|---|---|---|
| Serien-Umfrage in der Lehrveranstaltung | | | | | | | | |
| A | WS 02/03 | 40 | 39 | 97,50% | 4 | 2 | 5,1 min | 50 min |
| A | WS 03/04 | 46 | 45 | 97,83% | 4 | 2 | 6,2 min | 70 min |
| B | WS 02/03 | 57 | 55 | 96,49% | 4 | 2 | 4,4 min | 60 min |
| B | WS 03/04 | 49 | 45 | 91,84% | 5 | 2 | 6,7 min | 60 min |
| C | WS 02/03 | 18 | 16 | 88,89% | 3 | 1 | 4,7 min | 25 min |
| C | WS 03/04 | 45 | 43 | 95,56% | 5 | 1 | 6,4 min | 55 min |
| D | SS 03 | 30 | 29 | 96,67% | 4 | 2 | 5,5 min | 40 min |
| E | SS 03 | 27 | 25 | 92,59% | 4 | 2 | 4,8 min | 30 min |
| F | SS03 | 12 | 12 | 100,00% | 3 | 1 | 6,3 min | 25 min |
| Einzel-Umfrage außerhalb der Lehrveranstaltung mit Individual-Autorisierung | | | | | | | | |
| G | WS 02/03 | 16 | 14 | 87,50% | Evaluation über Internet. Alle Umfragen waren eine Woche online geschaltet. | | | |
| H | WS 03/04 | 125 | 87 | 69,60% | | | | |
| I | WS 03/04 | 298 | 137 | 45,97% | | | | |
| J | WS 03/04 | 227 | 88 | 38,77% | | | | |
| K | WS 03/04 | 524 | 211 | 40,27% | | | | |
| L | WS 03/04 | 468 | 199 | 42,52% | | | | |
| M | WS 03/04 | 264 | 80 | 30,30% | | | | |
| N | WS 03/04 | 130 | 57 | 43,85% | | | | |
| O | WS 03/04 | 360 | 97 | 26,94% | | | | |
| P | WS 03/04 | 216 | 71 | 32,87% | | | | |

Nicht zuletzt wegen des „Neuigkeitseffekts" dieser Befragungsart wurden diese Evaluationen sehr positiv aufgenommen. Bei den über Internet durchgeführten Einzel-Umfragen außerhalb der Lehrveranstaltung erkennt man, dass sehr unterschiedliche Rücklaufquoten erreicht wurden. Obwohl diese Rücklaufquoten im Vergleich zu den Werten, die man von der Marktforschung kennt, als in den meisten Fällen sehr hoch eingeschätzt werden können, hat sich eine Kommission, die den Testeinsatz der Evaluationssoftware auf Verbesserungsmöglichkeiten prüft, bereits Gedanken dazu gemacht, wie man den Rücklauf verbessern kann. Eine Empfehlung ist, möglichst kurze Evaluierungen zu konzipieren und evtl. getrennt nach Lehrveranstaltungsform abzufragen (z.B. die Evaluation einer Vorlesung nicht zusammen mit der zur Vorlesung gehörenden Übung stattfinden lassen). Zur Modifikation der Frageblöcke wurden aufgrund der gesammelten Evaluationsdaten auch Faktorenanalysen durchgeführt, um abschätzen zu können, welche Einzelfragen mit welchen Faktoren in Zusammenhang zu bringen sind und wie das Nichtberücksichtigen von Antworten einzelner Fragen die Faktorenstrukturen und ihren Erklärungsgehalt beeinflusst.

Mit der Durchführung der computergestützten Evaluationen wurde im Wintersemester 02/03 im Rahmen eigener Veranstaltungen begonnen, wobei zunächst mit

Serien-Umfragen in der Lehrveranstaltung gestartet wurde. Allerdings gehörte bereits damals auch eine Einzelumfrage außerhalb der Lehrveranstaltung (Vorlesung G) zum Probelauf. Im Wintersemester 03/04 wurden „größere" Lehrveranstaltungen (die Vorlesungen H bis P) des Grundstudiums evaluiert, wobei Einzelumfragen mit Individual-Autorisierung außerhalb der Lehrveranstaltung per Internet erfolgten. Einige Veranstaltungen (die Vorlesungen A, B, C) wurden im Wintersemester 03/04 in unveränderter Befragungsform wieder evaluiert, um Vergleichsdaten zu erhalten.

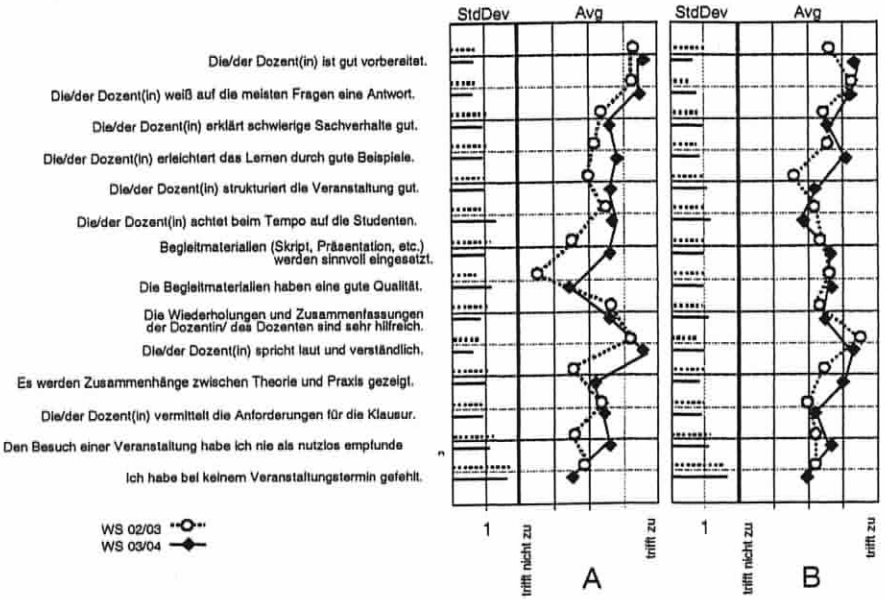
In Abbildung 4 werden die Ergebnisse des Fragenblocks „Leistungen der Dozentin/des Dozenten" für die Vorlesung A und B auf Basis der genannten Vergleichsdaten wiedergegeben. Gezeigt werden Durchschnitt (Avg) und Standardabweichung (StdDev) jeweils für die Umfragen aus WS 02/03 und WS 03/04. Die Antwortskala erstreckt sich über das Intervall [-2,+2] mit „trifft nicht zu" bis „trifft zu" als Extremausprägungen (siehe auch Abbildung 3).

Die Veranstaltungen A und B wurden von verschiedenen Dozenten angeboten. Die Diagramme wurden nebeneinander gestellt, um die Ergebnisse besser vergleichen zu können. Insgesamt erkennt man, dass sich der schwarze Linienzug (WS 03/04) gegenüber dem gepunkteten Linienzug (WS 02/03) fast überall in die „trifft zu"-Richtung verschoben hat. Auf positiv bewertete Aussagen soll hier nicht eingegangen werden. Es werden nur jeweils eher als nicht zutreffend bewertete Aussagen herausgegriffen, weil hieraus Erkenntnisse ableitbar sind, die in besonderem Maße der Qualitätsentwicklung dienen können. Positiv bewertete Aussagen können natürlich ebenfalls als qualitätsentwickelnde Anhaltspunkte dienen, werden hier aber nicht weiter angesprochen.

Teil A der Abbildung 4 zeigt Ergebnisse, durch die ein Dozent X während seiner Veranstaltung A evaluiert wurde. Interessant ist hier beispielsweise die Aussage „Die Begleitmaterialien haben eine gute Qualität". Die Begleitmaterialien dieser Vorlesung waren im WS 02/03 im Gegensatz zu anderen Vorlesungen des evaluierten Instituts überwiegend handschriftlich gefasst. Die Studierenden haben in der zweiten Umfrage im WS 03/04, als das Skript in wesentlichen Teilen von der handschriftlichen Vorlage in gedruckte Form umgestellt war, hier bereits deutlich positiver reagiert. Der Einsatz der Begleitmaterialien, der bei dieser Lehrveranstaltung im WS 02/03 ebenfalls bemängelt wurde, wird im WS 03/04 nun ebenfalls deutlich besser bewertet.

In Teil B der Abbildung 4 wurde ein anderer Dozent Y während seiner Veranstaltung B evaluiert. Im Vergleich zu den übrigen abgefragten Statements wurde lediglich die Aussage „Der Dozent strukturiert die Veranstaltung gut." im WS 02/03 als eher nicht zutreffend bewertet. In diesem Fall war es so, dass vorlesungsbegleitend das Skript überarbeitet wurde, was dazu führte, daß am Anfang der Vorlesung B im WS 02/03 nicht das komplette Skript verteilt wurde, sondern in zeitlichen Abständen jeweils nur einzelne überarbeitete Kapitel in der Reihenfolge der Fertigstellung und nicht in der Reihenfolge, in der das Skript und die Vorlesung aufgebaut war. Im WS 03/04 war das vollständige Skript bereits zu Beginn der Vorlesung B verfügbar, was den Dozenten veranlasste, „schneller" beim Durcharbeiten der Stoffinhalte vorzugehen, und sich in einer Verschlechterung bei der Bewertung der Aussage „Die/der Dozent(in) achtet beim Tempo auf die Studierenden" bemerkbar machte.

*Abbildung 4*: „Leistungen der Dozentin/des Dozenten" bei Vorlesung A und B (verschiedene Dozenten)




Diese von den Dozenten X, Y gegebenen Erklärungen sind aufgrund der Profilvergleiche in Abbildung 4 plausibel. Die Dozenten haben mit den Studierenden auch andere Unterschiede (bzw. Übereinstimmungen) in den Profilzügen zu Diskussionen genutzt, wobei sich im WS 03/04 zusätzlich angebotene Möglichkeiten zur Eingabe von Freitexten als sehr sinnvoll erwiesen haben, weil Häufungen entsprechender positiver/negativer Äußerungen den Erklärungsgehalt der Profilzüge verbessern konnten und außerdem Hinweise auf nicht abgefragte Aspekte erlaubten.

## 4.3 Aufwand für Evaluatoren und Administration

Das Anlegen einzelner Lehrveranstaltungsevaluationen im ETU NetEval System dauerte pro Veranstaltung ca. 5 Minuten. Für die praktische Durchführung der Serien-Umfragen in den Lehrveranstaltungen wurden Notebooks ausgeliehen und vorbereitet, was mit Hilfe des speziell dafür entwickelten ETU SurveyBrowser Programms einmalig ca. 5 Minuten pro Gerät beansprucht hat. Die Serien-Umfragen in den Lehrveranstaltungen wurden von ein oder zwei Personen beaufsichtigt. Die Vorgehensweise führte zu sehr guten Rücklaufquoten (vgl. Tabelle 2). ETU NetEval wurde von verschiedenen Dozenten der Universität Karlsruhe getestet. Dabei hat sich bestätigt, dass die Durchführung von Lehrveranstaltungsevaluationen mit ETU NetEval für die Evaluatoren kein spezielles Know-How erfordert und wenig Zeit beansprucht. Nicht zuletzt deshalb gab es von Seiten der Evaluatoren als Systemnutzer nur positive Reaktionen. Der Aufwand für die administrative Betreuung von ETU NetEval lag im üblichen Rahmen für die Wartung und Pflege eines Web-

Servers und bestand primär aus Routine-Aufgaben wie Backup und Installation von Sicherheitsupdates der zugrunde liegenden Systemsoftware (Linux, Apache, Tomcat, Oracle). Dieser Aufwand entstand unabhängig von der Nutzungsintensität von ETU NetEval und wird damit unter Ausnutzung von Skaleneffekten einen großflächigen Einsatz des Systems ermöglichen.

## 5. Ausblick

Ein System zur computergestützten Lehrveranstaltungsevaluation wird sich erst nach mehreren Runden eines praktischen Einsatzes einem „Design" nähern, welches von unterschiedlichen Zielsegmenten als gleichermaßen akzeptabel angesehen wird. Die Einbeziehung verschiedener Nutzergruppen in diesen Entwicklungsprozess ist hierbei von besonderer Wichtigkeit. Bisher hat sich bereits gezeigt, dass computergestützte Lehrveranstaltungsevaluationen für die Beteiligten an den diversen Testläufen mit einer Reihe von Vorteilen verbunden sind.

Der nächste Schritt wird deshalb die Gewinnung von weiteren Nutzern innerhalb der Universität Karlsruhe sein. Nachdem die Fakultät für Wirtschaftswissenschaften das System im Wintersemester 2003/04 getestet hat, wird die Fakultät für Mathematik das System zum Ende desselben Semesters ebenfalls in Probebetrieb nehmen. Wenn diese Tests erfolgreich verlaufen, soll das System für die nächsten Lehrevaluationen an den jeweiligen Fakultäten großflächig eingesetzt werden. Gespräche über die weitere Entwicklung des Systems und die Ausstattung mit zusätzlichen/geänderten Fragenblöcken haben begonnen. Wenn die Datenbasis durch wiederholte Evaluationen eines Veranstaltungstyps bzw. durch Evaluationen mehrerer verschiedener Veranstaltungen entsprechend umfangreich geworden ist, werden zusätzliche Auswertungsmöglichkeiten integriert und Anbindungen an professionelle Datenanalyse-Tools freigeschaltet werden.

Obwohl im vorliegenden Artikel lediglich eine computergestützte Evaluation herkömmlicher Lehrveranstaltungsformen (z.B. Vorlesungen/Übungen/Tutorien in einem Hörsaal) diskutiert wurde, bildet ein System wie ETU NetEval natürlich eine gute Voraussetzung für Evaluationen im Rahmen einer Notebook Universität und ist überall dort einsetzbar, wo mobile Lernumgebungen angewandt werden.

## 6. Literatur

ABET: Accreditation Board for Engineering and Technology, Inc. URL: http://www.abet.org, 24.06.03.

CLIX: Corporate Learning & Information eXchange, mc information multimedia communication AG, 66115 Saarbrücken. URL: http://www.clix.de, 04.06.03.

DUI: Dokumente & Informationen 1/1998: Evaluation – Sachstandsbericht zur Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen; Evaluation – State of the art report on quality assessment and quality development in German universities, Projekt Qualitätssicherung, Hochschulrektorenkonferenz (Hg.) (1998).

DUKATH: Drahtlose Universität Karlsruhe (TH). URL: http://www.uni-karlsruhe.de/~DUKATH/, 23.06.03.

EvaNet (1998): FAQ EvaNet. URL: http://evanet.his.de/evanet/knowhow/kh.grund/kh.faq.html, 12.06.03.

FORMGEN: pilodata GmbH, Peter-Henlein-Str. 4, D-82140 Olching. URL: http://www.formgen.de, 17.06.03.

HRG: Hochschulrahmengesetz, Begründung zu §6. URL: http://www.rewi.hu-berlin.de/stud/fsj/Hochschulrecht/bundesrecht/ehrg.html, 17.06.03.

INQUIRY: FH Ravensburg-Weingarten. URL: http://inquiry.fh-weingarten.de, 04.06.03.

Kromrey, H. (2001): Evaluation von Lehre und Studium – Anforderungen an Methodik und Design. In: Spiel, Christiane (Hg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck, S. 21-60.

LIQUIS: Johannes Kepler Universität Linz, SOWI – Fakultät. URL: http://liquis.ce.uni-linz.ac.at, 04.06.03.

Mitterauer, L. (1996): Evaluation im Hochschulbereich. In: Zeitschrift für Hochschuldidaktik, Nr. 1-2.

NUKATH: Notebook Universität Karlsruhe (TH). URL: http://www.nukath.uni-karlsruhe.de, 24.06.03.

Rindermann, H. (2001a): Die studentische Beurteilung von Lehrveranstaltungen – Forschung und Implikationen. In: Spiel, Christiane (Hg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck, S. 61-88.

Rindermann, H. (2001b): Lehrevaluation. Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Landau: Verlag für Empirische Pädagogik.

Schomann, T./Buhl, T./Rhinow, S. (2000): Handbuch Befragung (Version 1.1 – Jan. 2000). Universität Leipzig, Geschäftsstelle Evaluation.

Schratz, M. (1996): Qualitätssicherung in der universitären Lehre durch computergestützte Evaluation (Teaching Evaluation Assistant). In: Zeitschrift für Hochschuldidaktik, Nr. 1-2.

Souvignier, E./Gold, A. (2003): Fragebögen zur Lehrevaluation: Was können sie leisten? In: Zeitschrift für Evaluation, 2/2002, S. 265-280.

ULGE: Universität Leipzig, Geschäftsstelle Evaluation. URL: http://www.uni-leipzig.de/~eval/, 04.06.03.

VND: Verbund Norddeutscher Universitäten: Wir über uns. URL: http://www.zeva.uni-hannover.de/zeva.htm, 23.06.03.

ZEvA: Zentrale Evaluations- und Akkreditierungsagentur Hannover. URL: http://www.zeva.uni-hannover.de/zeva.htm, 18.06.03.