

Informationskompetenzveranstaltungen evaluieren

Qualitätskontrolle durch Befragung

Viola Hämmer, Eva Flammersfeld

1. Was versteht man unter Evaluation? Wozu wird evaluiert?

Kurse zur Vermittlung von Informationskompetenz sind seit geraumer Zeit aus dem Serviceangebot wissenschaftlicher Bibliotheken nicht mehr wegzudenken. Die Veranstaltungen zu Themen wie Bibliotheksbenutzung, Literaturrecherche oder zu urheberrechtlichen Fragen werden an Universitäten vermehrt in die Curricula der Fachbereiche eingebunden. Dies zeugt von der wachsenden Bedeutung der Bibliotheken bei der Wissensvermittlung. Die Rolle des Bibliothekars beinhaltet in einem wachsenden Anteil Lehrtätigkeit. Und damit stellt sich die Frage nach der Qualität bibliothekarischer Lehre. Diese lässt sich nicht einfach aus den inhaltlichen Kompetenzen ableiten, die Bibliothekare in vielen Fällen sicherlich zu Recht für sich reklamieren. Um dem Anspruch kompetent zu lehren, tatsächlich gerecht zu werden, ist es notwendig, die Qualität der Veranstaltungen regelmäßig zu prüfen. Es lohnt sich zu reflektieren, ob das Interesse der Teilnehmer geweckt wird, ob die angewandten Lehrmethoden zum Lernen motivieren und natürlich ob die Lerninhalte erfolgreich vermittelt werden.

Was ist unter Evaluation überhaupt zu verstehen? Clarke & Dawson prägten folgende Begriffsbestimmung: *„Evaluation is defined as a form of disciplined inquiry (...) that applies scientific procedures to the collection and analysis of information about the content, structure and outcome of programmes, projects and planned interventions“*¹. Im deutschsprachigen Raum hat sich u. a. Rindermann besonders mit studentischer Lehrevaluation beschäftigt. Er definiert Evaluation ganz ähnlich als einen *„Prozess der systematischen Beurteilung eines Programms oder Produkts“*². Beim Stichwort „Lehrevaluation“ liegt der Gedanke an studentische Bewertungen von Lehrveranstaltungen meist sehr nahe. In der (sozial-)wissenschaftlichen Diskussion wird allerdings ganz im Gegensatz dazu von manchen Autoren gar bezweifelt, ob es sich bei solchen Befragungen überhaupt um evaluative Maßnahmen handelt.³ Auf diese Diskussion wird hier nicht eingegangen. Der Artikel beschäftigt sich vorwiegend mit studentischer Beurteilung von Bibliotheks-

- 1 Clarke, Alan ; Dawson, Ruth: Evaluation Research. An Introduction to Principles, Methods and Practice. London [u.a.], 1999, S. 1.
- 2 Rindermann, Heiner: Die studentische Beurteilung von Lehrveranstaltungen, S. 13. In: Tests und Trends, 11 (1997), S. 12–53.
- 3 Vgl. z.B. Kromrey, Helmut: Studentische „Lehrevaluation“ oder (nur) „Teilnehmerbefragungen“ in Lehrveranstaltungen? Methodische Probleme bei der Bewertung von Lehrqualität. In: Gralki, Heinz O. ; Grün, Dieter ; Hecht, Heidemarie (Hg.): Evaluation schafft Autonomie. Perspektiven der Lehrbewertung an Hochschulen, Berlin, 1993, S. 43–51.

kursen. Die Erhebung und Analyse der Meinungen von Kursteilnehmern ist aus unserer Sicht eine leicht durchführbare und dennoch aussagekräftige Methode der Qualitätskontrolle. Wir schließen uns Rindermann & Amelang an, die konstatieren, „daß bei Beachtung einiger methodischer Sorgfaltsregeln von hinreichender Validität studentischer Evaluationen auszugehen ist“⁴. Erwähnenswert ist dabei auch, dass diese Evaluationsform im privatwirtschaftlichen Bereich überhaupt nicht mehr wegzudenken ist. Eigentlich alle Träger beruflicher Fortbildung nutzen dieses Instrument der Lehrbewertung, um ihr Angebot auf dem Laufenden zu halten und um konkurrenzfähig zu sein. Die Teilnehmer werden bei Befragungen als „Kunden“ betrachtet, deren Meinung eine wichtige Rolle spielt.

Im Prozess der Evaluation können Aspekte einer Veranstaltung, die von den Teilnehmern als negativ bewertet werden, ermittelt und im Folgeschritt analysiert, verbessert oder beseitigt werden. Im Gegenzug dazu können die als positiv beurteilten Aspekte beibehalten und verstärkt werden. Evaluative Maßnahmen dienen also der rationalen Entscheidungsfindung bei der Bewertung der Veranstaltung und haben letztlich deren Effizienzsteigerung zum Ziel. Eine kontinuierlich stattfindende Evaluation führt im besten Fall zu immer positiver ausfallenden Ergebnissen. Positive Evaluationsergebnisse können eine Veranstaltung in mehrererlei Hinsicht verbessern. Sie wirken sich auch auf die Motivation der Dozenten aus. Ein positives Feedback erzeugt „Freude am Gelingen“ und ein motivierter Dozent vermag seinen Lehrstoff erfolgreicher zu vermitteln als der Freudlose und Schlechtgelaunte.

2. Prinzipien der Lehrevaluation

Evaluationsmethoden sind vielfältig und müssen dem eigenen Bedarf angemessen ausgewählt und angepasst werden. Einige Prinzipien sollten allerdings bei jeder Methode beachtet werden.

Besonders wichtig ist es, im ersten Schritt die *Ziele*, die in der Lehrveranstaltung erreicht werden sollen, zu formulieren. Diese sind die Grundlage für die Evaluation. Erst wenn festgelegt wurde, was erreicht werden soll, kann auch ein Erfolg oder Misserfolg beobachtet werden. Dieser Punkt erscheint vielleicht trivial, in der Praxis stehen aber viele Evaluationsbemühungen genau an diesem Punkt

4 Rindermann, Heiner; Amelang, Manfred: Zur Validität der Beurteilungen von Lehrveranstaltungen durch Studierende. Manuskript für den 39. Kongreß der Deutschen Gesellschaft für Psychologie in Hamburg vom 25. bis 29. September 1994 in Hamburg. Heidelberg, 1994, S. 23. Zit. nach Schnell, Rainer; Kopp, Johannes: Theoretische und methodische Diskussion der Lehrevaluationsforschung und deren praktische Bedeutung. Projektbericht des Projekts, Fakultätsinterne Evaluation der Lehre: Die Weiterentwicklung des bisherigen Evaluationskonzepts. Konstanz, 2000, S. 45. <http://www.ub.uni-konstanz.de/kops/volltexte/2001/605/>

vor einem großen Problem. Häufig ist nicht klar, welche Dimensionen eines meist recht vielfältigen Bereiches bewertet werden sollen.

Ein Hauptziel jeder Veranstaltung ist die Vermittlung bestimmter Lerninhalte. Diese sollten zunächst spezifiziert werden: Geht es z.B. darum, den Teilnehmern einen groben Überblick über eine Thematik zu verschaffen? Das könnte etwa bei einem Überblick über die Internetressourcen eines Faches der Fall sein. Oder sollen die Teilnehmer nach der Veranstaltung die Basisfunktionen des UB-Katalogs sicher beherrschen? Das wäre ein ganz anderes Ziel, es müsste anders evaluiert werden. Neben der Inhaltsvermittlung kann es ein weiteres Ziel sein, das Interesse der Teilnehmer für die Thematik zu wecken oder eine bei den Teilnehmern möglicherweise vorhandene Scheu vor der Bibliothek abzubauen.

Nach der Spezifikation der Ziele geht es darum, eigene *Qualitätskriterien* für die Veranstaltung festzulegen. Bibliothekskurse können sehr unterschiedlich gestaltet werden. Es ist wichtig, sich selbst darüber klar zu werden, was eigentlich einen „guten Kurs“ ausmacht. Zu diesen Kriterien können z.B. eine klare Strukturierung der Veranstaltung, die Auswahl geeigneter Übungsbeispiele, angemessene technische und räumliche Verhältnisse oder gute rhetorische Fähigkeiten des Dozenten gehören.

Eingangs wurde auf methodische Sorgfaltsregeln verwiesen, die für die Zuverlässigkeit der Evaluation eingehalten werden sollten. Eine der wichtigsten Regeln betrifft die *Anonymität der Befragten*. Die Gewährleistung von Anonymität trägt dazu bei, Verfälschungen in den abgefragten Daten zu minimieren. Ein Hauptproblem bei fast jeder Befragung (und somit auch bei der Evaluation) sind „sozial erwünschte“ Antworten. Diese spiegeln nicht in erster Linie die „wahre“ Meinung der Teilnehmer zur Veranstaltung wider, sondern sie orientieren sich stark an Erwartungen und Normen. Bleiben Befragte nicht anonym, antworten sie tendenziell eher so, wie sie meinen, dass es der Dozent erwartet. Auch Antworten, die den Befragten in einem guten Licht erscheinen lassen, sind dann wahrscheinlich. Wer identifizierbar ist, wird nur selten zugeben, dass er trotz zweistündiger Veranstaltung nur wenig gelernt hat. Das ist bei der Erhebung der Daten zu berücksichtigen. In vielen Evaluationsfragebögen werden demographische Daten wie z.B. Alter, Geschlecht und Semesterzahl abgefragt. Bereits das kann in einigen Fällen die Anonymität der Befragten gefährden: Der Fragebogen des einzigen Mannes in einer Gruppe sonst weiblicher Germanistikstudentinnen ist bei Abfrage des Geschlechts identifizierbar. Im Zuge des Bologna-Prozesses werden bibliothekarische Lehrveranstaltungen vermehrt in die Lehrpläne der Studiengänge an Universitäten und Fachhochschulen eingebunden. Bibliothekare werden demzufolge nicht nur zu Lehrenden, sie werden häufig auch zu Prüfern. Und dann ist die Wahrung der Anonymität besonders wichtig. Denn gerade wenn ein Abhängigkeitsverhältnis zwischen Fragendem und Befragtem herrscht, muss man damit

rechnen, stark verzerrte Daten zu erhalten. Die Tendenz, im Sinne der sozialen Erwünschtheit zu antworten, kann niemals ganz ausgeschaltet werden. Viele Personen antworten auch dann in diese Richtung, wenn sie nicht beurteilt werden, vielleicht einfach deshalb, weil sie einem sympathischen Dozenten nichts Negatives sagen möchten. Man kann sich aber bei der Durchführung der Evaluation auf dieses Problem einstellen bzw. die Ergebnisse der Evaluation mit dem Wissen um dieses Problem bewerten.

Neben der Gewährleistung der Anonymität kann man durch die *Gestaltung der Instruktion* die Wahrscheinlichkeit „ehrlicher“ Antworten erhöhen. Bei der Einführung der Evaluation ist es wichtig, explizit darauf hinzuweisen, dass man an der tatsächlichen Meinung der Teilnehmerinnen und Teilnehmer interessiert ist. Dies kann mündlich oder schriftlich geschehen. Günstig ist in diesem Zusammenhang auch, deutlich zu machen, dass Kritik ernst genommen wird und Änderungen so weit wie möglich vorgenommen werden.

Für die Evaluation sollte *Veranstaltungszeit* reserviert werden, d. h. das jeweilige Verfahren sollte – sofern realisierbar – direkt im Kurs durchgeführt werden. Auch dies zeigt den Teilnehmern, dass die Evaluation ernst genommen wird, dass sie nicht nebenbei zu erledigen ist. Zudem müssten die Befragten, wenn sie die Fragebögen mit nach Hause nehmen, nochmals aktiv werden, um ihre Meinung tatsächlich abzugeben. Dies ist häufig nicht der Fall, die Rücklaufquoten sind bei dieser Befragungsart generell niedrig.⁵ Bei so geringen Response-Raten kann man die erhaltenen Daten kaum noch seriös verwenden. Es liegt dann lediglich die Meinung einer hochselektiven Stichprobe vor und nicht eine aussagekräftige Beurteilung.

Eine *zeitlich und organisatorisch* dem jeweiligen Rahmen der Veranstaltung angemessene Durchführbarkeit der angewandten Methode ist ebenso Voraussetzung für eine gelungene Evaluation. Bei einer einmaligen Veranstaltung sind andere Evaluationsmethoden auszuwählen als bei einem Kurs, der sich über ein ganzes Semester erstreckt. Die Evaluation von Einzelveranstaltungen muss natürlich knapper sein als die von mehrmaligen Kursen. Eine „formative“ Evaluation⁶, bei der in regelmäßigen Abständen über einen längeren Zeitraum (z.B. mehrmals im Semester) mit derselben Teilnehmergruppe evaluiert wird, kommt hier nicht in Frage. Diese wäre bei mehrmaligen Veranstaltungen prinzipiell möglich, wird aber aufgrund des hohen Aufwands nur sehr selten praktiziert. Die einmalige Evaluation am Ende einer Veranstaltung nennt Scriven „summativ“, sie ist die am weitesten verbreitete Form der Lehrevaluation. Als sehr günstig bei längeren Kursen hat

5 Vgl. z.B. Schnell, Rainer ; Hill, Paul ; Esser, Elke: Methoden der empirischen Sozialforschung. München, 2005.

6 Begriff nach Scriven, Michael: The Methodology of Evaluation. In: Gredler, Margaret E.: Program Evaluation. New Jersey, 1996, S. 16.

sich allerdings eine Zwischenevaluation erwiesen, die im Sinne einer formativen Herangehensweise eine noch laufende Veranstaltung überprüft.⁷ Die Kursteilnehmer werden dabei nach der Hälfte der Veranstaltung um ihre Meinung gebeten. Sie haben dann ein Eigeninteresse an einer gewissenhaften Evaluation, da sich ihre Beurteilungen noch für sie selbst positiv auswirken können. Die Platzierung der Evaluation am Ende einer Veranstaltung dagegen führt niemals dazu, dass die Teilnehmer aus Eigeninteresse mitmachen. Zudem besteht hier noch ein weiterer Nachteil: Es ist selbst bei Pflichtveranstaltungen mit einer positiven Selektion und entsprechenden Verzerrungen zu rechnen.⁸ Häufig geraten bei diesem Verfahren die besonders unzufriedenen Studierenden gar nicht mehr in die Evaluations-situation, da sie möglicherweise gar nicht mehr anwesend sind.

Generell ist es bei freiwilligen Befragungen von entscheidender Bedeutung, die *Befragten zur Mitarbeit zu motivieren*. Diese Bereitschaft hängt wesentlich von der zu investierenden Zeit, vom Layout des Fragebogens (bei schriftlichen Befragungen) und der Art der Einstiegsfragen ab. Die ersten Fragen sollten weder zu komplex noch zu eintönig sein. Beides wirkt sich demotivierend auf den Befragten aus und kann zum Abbruch oder zum beliebigen Antworten führen. Das Layout sollte klar, übersichtlich und ansprechend sein (Hippler, 1988 in Rammstedt). Besonders die Knappheit ist ein wichtiges Prinzip. Wird die Evaluation zu lang und von den Teilnehmern als unangemessen erlebt, kann diese Maßnahme der Qualitätssicherung eine bis dahin gelungene Veranstaltung kaputt machen und somit ihrer ursprünglichen Intention entgegenlaufen.

3. Methoden der Evaluation

Im Folgenden sollen einige Evaluationsmethoden vorgestellt werden. Welche Methode für die eigene Veranstaltung am besten geeignet ist, sollte im Einzelfall entschieden werden.

Die *Punkt-Abfrage* ermittelt mit relativ einfachen Mitteln ein Stimmungsbild zu einzelnen Fragestellungen. Dazu werden zunächst eine oder mehrere Skalen auf (ein) Plakat(e) aufgezeichnet. Diese Skalen werden jeweils mit zwei gegensätzlichen Aussagen wie z.B. „*Mit der Veranstaltung war ich sehr unzufrieden*“ auf der linken Seite und „*Mit der Veranstaltung war ich sehr zufrieden*“ auf der rechten Seite beschriftet (oder „*Heute ging mir alles viel zu schnell*“ und „*Heute war das Tempo*

7 Vgl. Rindermann, Heiner ; Amelang, Manfred: Das Heidelberger Inventar zur Lehrveranstaltungs-Evaluation (HILVE) (Fragebögen, Kurzanleitung und Manual). Heidelberg, 1994, S. 10 f.

8 Vgl. z.B. Daniel, Hans-Dieter: Studentische Beurteilung von Lehrveranstaltungen – Anlage, Durchführung und Ergebnisse eines Modellprojekts an der Universität Mannheim. In: Hochschulrektorenkonferenz (Hg.): Evaluation und Qualitätssicherung an den Hochschulen in Deutschland – Stand und Perspektiven. Bonn, 1998, S. 79–104.

ziemlich schleppend“). Es können auch zwei Skalen miteinander kombiniert werden. Dann werden die beiden Skalen wie in einem Koordinatensystem aufgetragen. Die Teilnehmer bekommen selbstklebende Farbpunkte und kleben einen Farbpunkt auf den für sie zutreffenden Skalenabschnitt bzw. an die zutreffende Stelle des Koordinatensystems. Bei Bedarf kann das Ergebnis unmittelbar mit den Teilnehmern besprochen werden, so dass sich zusätzlich zur graphisch dargestellten Evaluation ein mündliches Feedback ergibt.

Die Punkt-Abfrage ist ein geeignetes Mittel, um die Tendenz bezüglich der Stimmung unter den Teilnehmern zu einzelnen Aspekten der Veranstaltung zu ermitteln. Für die Studenten ist diese Evaluation schnell und einfach zu erledigen, meistens macht sie den Befragten sogar Spaß. Die Vorbereitung (Plakate erstellen, Klebepunkte bereitlegen) ist allerdings aufwändiger als das bloße Kopieren eines vorbereiteten Fragebogens. Und letztlich ist der Informationsgewinn nur auf eine bzw. wenige Fragestellung/en beschränkt. Für eine umfassende, auf verschiedene Teilaspekte einer Veranstaltung bezogene Evaluation ist die Methode weniger geeignet. Ein weiterer Nachteil der Methode kann in der Art der Durchführung liegen: Wenn die Teilnehmer ihre Stimme öffentlich abgeben, bedeutet das, dass keine echte Anonymität gewährleistet ist. Zwar sind die Punkte nicht direkt einer Person zuzuordnen, aber die Person wird beobachtet, während sie ihren Punkt klebt. Dem kann man entgegenwirken, indem das Plakat auf einer Flipchart befestigt und diese so gedreht wird, dass die Teilnehmer beim Aufkleben der Punkte nicht beobachtet werden können. Ein weiterer Nachteil liegt in Gruppeneffekten, die die Bewertungen beeinflussen: Hat der erste Student seinen Punkt an eine bestimmte Stelle geklebt, beeinflusst dies die nachfolgenden Befragten mehr oder weniger stark. Viele Personen orientieren sich in ihrem Urteil nun auch an der schon vorhandenen Einschätzung. Bei Evaluationsmethoden, die eine solche offene Beurteilung enthalten, ist eine geringere Streuung der Daten zu erwarten. Die Urteile konzentrieren sich aber nicht notwendigerweise auf die „wahre“ durchschnittliche Einschätzung, sondern auf die Meinung eines Kursteilnehmers, nämlich desjenigen, der als erster sein Urteil abgibt.

Die *Blitzlicht-Methode* besteht ausschließlich aus mündlichem Feedback. Alle Teilnehmer werden am Ende der Veranstaltung dazu aufgefordert, in kurzer Form ihre Meinung zur Veranstaltung zu äußern. Man kann das Blitzlicht begrenzen (z.B. „Ein charakterisierender Satz zur Veranstaltung“) oder jeden so viel oder so wenig sagen lassen, wie er möchte. Die Beiträge bleiben unkommentiert.

Um eine schnelle Durchführung zu gewährleisten, sollte ein Blitzlicht nur in kleinen Gruppen Anwendung finden. Ein Blitzlicht kann nicht nur am Ende, sondern auch während einer Veranstaltung durchgeführt werden. So können die gewonnenen Informationen helfen, spontan zu reagieren und den weiteren Verlauf der Veranstaltung den Wünschen der Teilnehmer anzupassen.

Auch bei der Blitzlicht-Methode kann die Schwierigkeit darin bestehen, dass Befragte ihre Meinung den Vorrednern anpassen. Weiterhin ist keine Anonymität gegeben. Für eine solch offene Bewertung scheint eine gewisse Bekanntheit der Teilnehmer untereinander wichtig zu sein. In völlig unvertrauten Gruppen (z.B. kennen sich bei einmaligen Veranstaltungen oft weder die Teilnehmer, noch kennen die Studierenden den Dozenten) erbringt das Blitzlicht möglicherweise keine besonders aussagekräftigen Ergebnisse. Es könnten Vorbehalte bestehen, ehrliche Kritik zu üben und sich negativ über die Veranstaltung oder den Dozenten zu äußern. Falls sich dennoch ein Teilnehmer dazu entschließt, seine Meinung offen auszusprechen, geschieht das unmittelbar gegenüber dem Dozenten. Dieser hat dann wiederum die Schwierigkeit, möglicherweise entstandene Irritationen verbergen zu müssen, denn die Beiträge sollen ja unkommentiert bleiben.

Zur Evaluation und Optimierung von Veranstaltungen kann es sinnvoll sein, Beobachtungen unbeteiligter und damit objektiverer Personen zu Hilfe zu nehmen. Das *Peer-Feedback* sieht vor, eine außenstehende Person an der Veranstaltung teilnehmen und den Ablauf von ihr beobachten zu lassen. Meist ist es sinnvoll, dabei einen Auswertungsbogen zu Hilfe zu nehmen, in den der Beobachter seine Bemerkungen direkt einträgt.

Es muss geklärt werden, wer diese Beobachtungen machen kann. Ein Vertrauensverhältnis zwischen Beurteilendem und Beurteiltem ist eine Voraussetzung dieser Methode. Hierarchische Über- oder Unterordnungen sollten vermieden werden. Das Verfahren ergibt mit Sicherheit interessante Ergebnisse, da die Bewertungsperspektive eine ganz andere als die der Teilnehmer ist. Das *Peer-Feedback* ist als eine „*Expertenevaluation*“ anzusehen, da die beurteilende Person meist ebenfalls Bibliothekar ist und selbst Veranstaltungen durchführt. Allerdings muss bedacht werden, dass es sich bei dieser Bewertung um das subjektive Empfinden einer einzelnen beobachtenden Person handelt. Die Meinung der Teilnehmer wird außer Acht gelassen.

Das *schriftliche Kurzfeedback* fußt dagegen wieder auf der Meinung der Teilnehmer. Ein solches Feedback am Ende jeder Veranstaltung durchzuführen, ist unkompliziert und ergibt für den Dozenten schnell eine aussagekräftige Rückmeldung. Es handelt sich um einen Fragebogen, auf dem nur wenige, eher allgemein gestellte offene Fragen zu beantworten sind. Möglich sind Fragen wie z.B. „*Was hat Ihnen an der Lehrveranstaltung besonders gefallen?*“, „*Was hat Sie an der Lehrveranstaltung besonders gestört?*“, „*Was sollte beim nächsten Mal besser gemacht werden?*“, „*Was haben Sie gelernt?*“ und „*Was hat Ihnen gefehlt?*“. Jeder Teilnehmer kann seine Antworten nach Belieben formulieren.

Der Informationsgehalt bei einem schriftlichen Kurzfeedback ist als sehr hoch anzusehen. Je offener die Fragen gestellt werden, desto größer ist der Spielraum, der dem Teilnehmer für seine Antworten bleibt. Die Methode ist besonders geeignet,

wenn eine Veranstaltung zum ersten Mal angeboten wird. Die Antwortmöglichkeiten werden hier – ganz im Gegensatz zum Fragebogen mit sog. geschlossenen Fragen – nicht schon von vornherein auf Einzelaspekte eingegrenzt. Man kann so auch überraschende Antworten erhalten, nach denen man von selbst nicht gefragt hätte. Wenn man allerdings begründet weiß, was genau man abfragen will, bieten sich konkretere Fragen mit Antwortalternativen an.

Das Kurzfeedback ist relativ schwer statistisch auszuwerten, da die Antworten keinem vorgegebenen Schema folgen. Damit soll aber nicht gesagt sein, dass diese Daten gar nicht auswertbar wären. Gerade bei größeren Stichproben ist das Kurzfeedback eine Methode, die durchaus sehr stichhaltige Ergebnisse liefern kann. Die Auswertung dieser freien Antworten ist allerdings sehr aufwendig.

4. Evaluation mittels Fragebogen

Die gängigste Evaluationsmethode ist der Einsatz eines klassischen Fragebogens. Die UB Würzburg verwendet für die Evaluation ihrer Katalogeinführungen einen Fragebogen, bei dem geschlossene Fragen mit vorgegebenen Antwortalternativen und offene Fragen miteinander kombiniert werden. Die Daten eines Fragebogens mit Auswahlmöglichkeiten sind statistisch leichter auswertbar und besser miteinander zu vergleichen als die der offenen Fragen. Außerdem können auf einfache Weise unterschiedliche Aspekte gezielt abgefragt werden. Welche Aspekte das sein sollen, sollte bei der Entwicklung des Fragebogens genau bedacht werden.

Anregungen zur Gestaltung des Fragebogens bieten bereits vorhandene Verfahren der Lehrevaluation. Das bekannteste Instrument im deutschsprachigen Raum ist das Heidelberger Inventar zu Lehrveranstaltungsevaluation (HILVE) von Rindermann und Amelang.⁹ Das Verfahren besteht zum größten Teil aus standardisierten Fragen mit vorgegebenen Antwortalternativen. Insgesamt beinhaltet der Fragebogen 43 Items (=Fragen) plus demographische Angaben und offene Fragen und ist damit sicherlich zu ausführlich für die Beurteilung bibliothekarischer Veranstaltungen. Der Fragebogen soll dennoch kurz vorgestellt werden, da Dimensionen und Items methodisch nach den Standards der empirischen Sozialforschung konstruiert wurden. Die Items sind insgesamt 14 Dimensionen zugeordnet. Beispiele dieser Dimensionen sind:

- Struktur der Veranstaltung, z.B. *„Der inhaltliche Aufbau der Veranstaltung ist logisch nachvollziehbar“*
- Lehrkompetenz des Dozenten, z.B. *„Der Dozent kann Kompliziertes verständlich machen“*

9 Vgl. Fußnote 7.

- Klima der Veranstaltung, z.B. *„Der/Die Dozent/in ist im Umgang mit den Studierenden freundlich und aufgeschlossen“*
- Diskussion, z.B. *„Es finden ausreichend Diskussionen statt“*.

Weitere Dimensionen sind Dozentenengagement, Interessantheit der Veranstaltung, Über- oder Unterforderung der Teilnehmer oder Interesse der Studierenden am Thema. Zusätzlich wird eine allgemeine Beurteilung der Veranstaltung auf einer Notenskala von 1 (= sehr gut) bis 6 (= ungenügend) erfragt. Zwei Beispiele der offenen Fragen sind: *„Was ist besonders gut an der Veranstaltung?“*, *„Was ist schlecht?/Verbesserungsvorschläge“*.

Es geht in diesem Fragebogen also nicht nur um die Inhalte der Veranstaltung bzw. darum, ob die Lernziele erreicht wurden. Dies ist bei einem solchen „Universalfragebogen“ auch nur bis zu einem gewissen Grad möglich – denn die Inhalte bzw. die Lernziele beziehen sich ja immer auf eine konkrete Veranstaltung. Und gerade in diesem Punkt werden studentische Selbsteinschätzungen stark kritisiert. Die Aussage darüber, ob man etwas gelernt hat und der tatsächliche Lernerfolg müssen keineswegs deckungsgleich sein. Im Fragebogen kann man aber nur Selbsteinschätzungen erfassen. Um den Lernerfolg zu erfassen, müsste man testende Verfahren einsetzen, die aber (zumindest im deutschsprachigen Raum) extrem selten sind angesichts des Aufwands, den man dabei betreiben muss.¹⁰ Die Mehrdimensionalität von HILVE zeigt aber deutlich, dass auch Aspekte wie das Klima oder Dozentenverhalten eine wichtige Rolle bei der Beurteilung der Qualität einer Veranstaltung spielen.

In den Fragebogen können also zum einen Fragen eingearbeitet werden, welche die Realisierung der Lernziele abfragen. Dabei sollte man sich aber bewusst sein, dass man eine Einschätzung erhält und nicht eine Aussage darüber, ob das jeweilige Lernziel tatsächlich erreicht wurde. Außerdem kann abgefragt werden, inwieweit die Methodik der Veranstaltung (Vortrag, Learning by doing, Gruppenarbeit etc.) von den Teilnehmern angenommen wird. Auch die Gestaltung der Veranstaltung durch den Dozenten selbst sollte stets in Frage gestellt werden. Ein und dasselbe Konzept eines Kurses wird von unterschiedlichen Dozenten unterschiedlich verwirklicht. Der eine trägt strukturierter vor als der andere, manche haben ein besonders gutes Gespür dafür, was bei den Teilnehmern gerade vorgeht und eine dritte Dozentengruppe zeichnet sich besonders dadurch aus, die Teilnehmer zum Mitdenken und Mitmachen anzuhalten (und dies meint nicht nur das „Einsetzen einer Methode“, z.B. Learning by doing – es ist vielmehr eine Dozentenvariable). In nahezu allen „professionellen“ Evaluationsverfahren finden sich vielfältige Fragen

10 Etwa der Vergleich von Literaturlisten vor und nach den Bibliothekskursen. Vgl. z.B. Emmons, Mark ; Martin, Wanda: Engaging Conversation: Evaluating the Contribution of Library Instruction to the Quality of Student Research. *College & Research Libraries*, 63 (2002), H. 6, S. 545–60.

über den jeweiligen Dozenten und dessen Vorgehens- und Verhaltensweisen. Dies scheint auch für Bibliotheksveranstaltungen sinnvoll. Gerade an dieser Stelle sollten die Mitarbeiter aber frühzeitig in die Fragebogenkonstruktion einbezogen werden. Die Hintergründe sollten deutlich gemacht werden und die Ergebnisse zu diesen Fragen sollten nur der bewerteten Person selbst zugänglich sein.

Die Erfassung von Vorwissen und Interesse kann erwogen werden, da diese Variablen zwar in keinem direkten Zusammenhang zur Veranstaltung stehen, die Bewertung durch die Teilnehmer aber durchaus beeinflussen können. Viele Studien weisen vor allem einen Zusammenhang zwischen dem (Vor-)Interesse der Teilnehmer an der Thematik und den Lehrveranstaltungsbeurteilungen nach.¹¹ Teilnehmer mit hohem Interesse am Thema bewerten Veranstaltungen tendenziell besser als Uninteressierte. Spiel empfiehlt daher, in jedem Fall das Interesse der Teilnehmer an der Lehrveranstaltung zu kontrollieren. Einige Fragebögen beinhalten auch Items, die nach dem Grund des Veranstaltungsbesuchs fragen (Pflichtveranstaltung, persönliches Interesse, Empfehlung von Kommilitonen etc.).

Zur Gestaltung der einzelnen Items eines Fragebogens haben zahlreiche Autoren Regeln formuliert¹²: Die Formulierung der Fragen soll sich an „*der Alltagssprache (Umgangssprache) des durchschnittlichen Mitglieds der Zielpopulation*“¹³ orientieren. Das heißt, die Fragen sind für Schüler anders zu formulieren als für wissenschaftliche Mitarbeiter. Fragen sollten formal möglichst einfach, kurz und ohne Fachtermini gestellt werden. Die Fragestellungen müssen eindeutig sein. So kann z.B. die Frage „*Konnten Sie den Dozenten gut verstehen?*“ sowohl auf das akustische als auch auf das inhaltliche Verständnis abzielen. Auch die Verwendung mehrerer Aussagen in einer Frage (z.B. „*Die Präsentation war übersichtlich und interessant.*“) sollten vermieden werden. Doppelte Negationen (z.B. „*Der Dozent war nicht unfreundlich.*“) klingen kompliziert und können zu Missverständnissen führen. Auch Suggestivfragen („*Fanden Sie den Dozenten nicht auch sehr hilfsbereit?*“) und Fragen nach zu persönlichen Angaben sollten vermieden werden.

Werden zusätzlich offene Fragen in den Bogen integriert, besteht häufig die Problematik, dass diese von den Teilnehmern nur ungern, sehr kurz oder gar nicht beantwortet werden. Dies ist deswegen ungünstig, da die Antworten auf die of-

11 Vgl. Spiel, Christiane ; Gössler, P. Martin: Zum Einfluß von Biasvariablen auf die Bewertung universitärer Lehre durch Studierende. In: Zeitschrift für Pädagogische Psychologie, 14 (2000), H. 1, S. 38–47.

12 Vgl. z.B. Fisseni, Hermann-Josef: Lehrbuch der psychologischen Diagnostik. Göttingen, 2004, S. 206.

13 Tränkle, Ulrich: Fragebogenkonstruktion. In: Feger, Hubert ; Bredenkamp, Jürgen (Hg.): Datenerhebung. Enzyklopädie der Psychologie. Themenbereich B: Methodologie und Methoden. Göttingen, 1983, S. 262.

fenen Fragen für die Dozenten häufig aussagekräftiger sind. Diese Fragen werden häufiger beantwortet, wenn sie entsprechend formuliert werden. Möglich sind Formulierungen wie „Was sollte unbedingt an der Veranstaltung geändert werden?/ Was sollte unbedingt beibehalten werden?“ oder „Was hat Ihnen überhaupt nicht/ besonders gefallen?“. Eine einleitende Wendung kann die hohe Bedeutung der offenen Fragen noch einmal verstärken. Formulierungen wie „Überlegen Sie noch einmal ganz genau ...“, „Blicken Sie bitte noch einmal auf die Veranstaltung zurück ...“ oder „Für uns ist es besonders wichtig, dass Sie jetzt noch einmal genau überlegen ...“ erschweren es dem Teilnehmer, die Frage zu übergehen. Es kann auch erwogen werden, die offenen Fragen gleich an den Anfang des Fragebogens zu stellen.

Nach der Formulierung der Items müssen die Antwortskalen gestaltet werden. Meistens ist es günstig, dem Befragten die Möglichkeit zu geben, seine Antwort zu skalieren. Häufig wird empfohlen, nicht mehr als sieben Möglichkeiten für die Beantwortung zur Auswahl zu stellen, denn bei einer zu hohen Anzahl an Antwortalternativen besteht die Gefahr des beliebigen Ankreuzens. Zu wenige Alternativen können aber auch ungünstig sein. Bei drei Antwortalternativen bleiben nur zwei Extreme und der Mittelpunkt, z.B. sehr gut – neutral – nicht gut. Differenzierte Urteile lassen sich auf dieser Skala nicht unterbringen.

Bei den Antwortalternativen können entweder sämtliche Abstufungen verbal bezeichnet werden oder lediglich die Endpunkte der Ratingskalen. Bei nicht vollständiger verbaler Bezeichnung der Abstufungen empfehlen manche Autoren, diese zumindest mit Zahlen zu kennzeichnen.

Es stellt sich weiterhin die Frage, ob eine gerade oder eine ungerade Anzahl von Antworten vorgegeben werden soll. Die ungerade Anzahl bietet eine Mittelkategorie (für eine neutrale Antwort), die geradzahlige Variante folgt dem „forced choice-Paradigma“: Der Befragte muss sich entscheiden, ob er in die eine oder andere Richtung bewertet. Die Verwendung geradzahliger oder ungeradzahliger Abstufungen wird in der Literatur umfassend diskutiert.¹⁴ Für beide Möglichkeiten lassen sich Argumente anführen. Fasst man jedoch die Studien zusammen, deutet die Befundlage darauf hin, dass eine Ratingskala ohne Mittelpunkt dazu führt, dass solche Fragen vermehrt nicht beantwortet werden.¹⁵

14 Vgl. Mummendey, Hans-Dieter ; Grau, Ina: Die Fragebogen-Methode. Göttingen, 2008, S. 75 ff.

15 Vgl. Rammstedt, Beatrice: Fragebogen. In: Petermann, Franz ; Eid, Michael (Hg.): Handbuch der psychologischen Diagnostik. Göttingen, 2006, S. 109–117.

5. Der Würzburger Fragebogen

An der Universitätsbibliothek Würzburg hat sich ein breites Schulungsangebot für Hochschulangehörige, Schüler und andere externe Bibliotheksbenutzer etabliert. Einen großen Teil der Veranstaltungen nehmen in die Curricula eingebundene Kurse zur Bibliotheksbenutzung und Einführung in den Katalog ein. Diese Veranstaltung soll nun kurz geschildert werden, um danach den dazugehörigen Evaluationsfragebogen vorzustellen. Die Dozenten (bei Gruppen ab ca. 15 Teilnehmern wird der Kurs in der Regel von zwei Diplom-Bibliothekaren gehalten) geben zunächst eine Powerpoint-unterstützte Präsentation über den Aufbau der UB und deren Serviceangebote sowie einen ersten Überblick über die Homepage und den Katalog der UB. Anschließend bearbeiten die Studenten in einer ca. 45-minütigen Übungsphase selbständig oder in Gruppen Aufgaben zur Literaturrecherche am PC. Je nach Erklärungsbedarf werden die Aufgaben im Anschluss für alle vorgeführt und besprochen. Nach einem gemeinsamen Rundgang durchs Haus wird die Veranstaltung durch die Evaluation mittels eines einseitigen Fragebogens abgeschlossen. Dieser wird im Schulungsraum ausgeteilt. Die Dozenten bitten darum, ihn gleich auszufüllen und zurückzugeben. Der Würzburger Fragebogen erklärt in einer Überschrift und einem einleitenden Satz Zweck und Bedeutung der Evaluation:

Ihre Meinung ist uns wichtig!

Die Mitarbeiter/innen der Universitätsbibliothek sind ständig bemüht, ihre Einführungsveranstaltungen zu verbessern. Sie helfen uns dabei, wenn Sie uns diesen Fragebogen ausgefüllt zurückgeben. Wir freuen uns über Ihre Anregungen!

Grundlegende Lernziele der Veranstaltung sind, die Studierenden mit der Katalogrecherche vertraut zu machen und ihnen eine erste räumliche Orientierung im Gebäude zu vermitteln. Ob diese Lernziele aus der Sicht der Studierenden erreicht wurden, wird mit den ersten beiden Fragen überprüft:

1. Ich kenne mich jetzt aus ...	sehr gut	befriedigend	unzureichend
im Katalog der UB	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
im Gebäude der UB	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Ich kann jetzt ...	trifft zu		trifft nicht zu
ein Buch in der UB finden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
eine Zeitschrift in der UB finden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Literatur zu einem Thema finden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Nun folgen zwei Fragen zur Methodik der Veranstaltung. Es soll geprüft werden, ob die „Learning-by-doing-Methode“ von den Studierenden als positiv empfunden und eine Besprechung der Übungsaufgaben als sinnvoll bewertet werden.

	trifft zu			trifft nicht zu	
3. Das Lernen anhand der Übungsaufgaben hat mir gefallen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Eine abschließende Besprechung der Übungsaufgaben halte ich für sinnvoll.	ja <input type="checkbox"/>	nein <input type="checkbox"/>			

Der Dozent kann durch seine Person, seine didaktischen Fähigkeiten und seine Art und Weise der Präsentation und Gestaltung der Veranstaltung deren Bewertung enorm beeinflussen. Daher wird mit einer Frage auch die Zufriedenheit der Teilnehmer mit der Lehrkraft erfragt. In einem zusätzlichen Bemerkungsfeld hat der Teilnehmer die Möglichkeit, Erläuterungen hinzuzufügen. Das Bemerkungsfeld wurde deshalb ergänzt, da die Beurteilung des Dozenten in dieser einen Frage nur sehr eingeschränkt möglich ist. Die Teilnehmer haben so die Möglichkeit, auf spezielle einzelne Aspekte einzugehen.

	trifft zu			trifft nicht zu	
5. Mit den Dozent/innen war ich zufrieden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bemerkung:					

Die Wichtigkeit der in zwei Spalten stehenden offenen Fragen wird mit einem einleitenden Satz besonders betont. Die Formulierungen sollen die Teilnehmer dazu anregen, die Veranstaltung noch einmal als Ganzes zu betrachten und zu bewerten.

6. Blicken Sie bitte noch einmal auf die Veranstaltung zurück ...

Was hat Ihnen besonders gefallen?

Und was hat Sie am meisten gestört?



Ein Dank an die Teilnehmer für die Teilnahme an der Evaluation schließt den Fragebogen ab.

Wir bedanken uns für Ihre Rückmeldung!

Bei der Konzeption des Würzburger Fragebogens war die Zielgruppenorientierung ein wichtiges Prinzip. Die Teilnehmer sind überwiegend Studienanfänger im Alter von ca. 20 Jahren. Es wurde versucht, sich bei der Formulierung der Fragen dem Sprachgebrauch dieser Altersklasse anzunähern. Ein Fragebogen, der auf wissenschaftliche Mitarbeiter oder Professoren ausgerichtet ist, würde sprachlich anders gestaltet werden.

Die Kürze und Übersichtlichkeit regt zur Teilnahme an der Evaluation an und trägt zu einer hohen Rücklaufquote der Fragebögen bei. So konnte die UB Würzburg z.B. im Wintersemester 2007/08 in den Katalogeinführungen eine Rücklaufquote von 97% feststellen.

Der Würzburger Fragebogen ist als ein Beispiel eines Evaluationsfragebogens zu verstehen. Er wurde an die konkrete Veranstaltung und die Rahmenbedingungen vor Ort angepasst.

6. Auswertung der Ergebnisse und Konsequenzen der Evaluation

Unabhängig von der Methode, die zur Evaluierung verwendet wurde, kommt dem Dozenten nun die Aufgabe zu, die Daten auszuwerten und die Ergebnisse für die Optimierung seiner Veranstaltung zu nutzen. Über die Veränderungsprozesse im Anschluss an Evaluationen gibt es allerdings zahlreiche kritische Stimmen. So konstatieren Schnell und Kopp: *„Die Vorstellung, dass allein die Information über die geringe Qualität der Lehrveranstaltung genügen könnte, in folgenden Seminaren und Vorlesungen besser zu werden, entspricht kaum dem Stand der Forschung im Bereich der Handlungstheorien.“*¹⁶ Sohr spricht sogar von einer *„totalen Konsequenzlosigkeit“*¹⁷ vieler Evaluationsbemühungen. Auch an dieser Stelle des Evaluationsprozesses können aus unserer Sicht aber Bedingungen gesetzt werden, die begünstigen, dass eine solche Konsequenzlosigkeit nicht eintritt.

An erster Stelle jeder Auswertung steht dabei das schlichte „zur Kenntnis nehmen“ der Ergebnisse, seien es Stimmungsbilder einer Punkt-Abfrage, ausgesprochene Bemerkungen bei einem Blitzlicht oder niedergeschriebene Ausführungen

¹⁶ Schnell/Kopp, S. 37. Vgl. Fußnote 4.

¹⁷ Sohr, Sven: Studentische Evaluation: Ein geistiger Schildbürgerstreich?, S. 170. In: Berendt, Brigitte; Stry, Joachim (Hg.): Evaluation zur Verbesserung der Lehre und weitere Maßnahmen. Weinheim, 1993, S. 168–172.

auf einem Fragebogen. Wichtig ist dabei, Kritik zuzulassen und anzunehmen. Soll die Evaluation wirklich helfen, Schulungsangebote zu verbessern, nützt es nichts, Ausreden für ein schlechtes Ergebnis zu suchen oder dieses nicht ernst zu nehmen. Genauso falsch wäre es, eine freundliche Äußerung, ein Wort des Dankes oder ein großes Lob nicht anzunehmen, anstatt sich darüber zu freuen und als Motivation zu nutzen.

Über das persönliche Annehmen der Ergebnisse hinausgehend, sollten die erhobenen Daten gesammelt und ausgewertet werden. Das ergibt bei Methoden wie dem schriftlichen Kurzfeedback oder einem ausführlichen Fragebogen mehr Sinn als bei dem Plakat der Punkt-Abfrage. Auch mündliche Feedbacks wie die Blitzlicht-Methode sind für eine dauerhafte, statistische Auswertung weniger geeignet.

Die Erfassung der Daten kann je nach Geschmack und Aufwand, der betrieben werden soll, in Statistik-Programmen wie SPSS, in Excel-Tabellen oder bei vornehmlich offenen Fragen in einfache Worddateien übernommen werden. Die UB Würzburg erfasst die Ergebnisse der Evaluationsbögen in Excel-Tabellen, mit denen sich die Daten mit einfachen Mitteln auch graphisch gut darstellen lassen. Über die Bildung von Mittelwerten lassen sich die Ergebnisse leicht miteinander vergleichen und über einen längeren Zeitraum beobachten. Neben den Mittelwerten kann die Streuung der einzelnen Beurteilungen zur Interpretation des Gesamtbildes herangezogen werden. Dadurch wird die Auswertung aufwendiger und komplexer, in manchen Fällen aber auch aussagekräftiger. Dies gilt besonders bei einer großen Meinungs- und Urteilsvielfalt unter den Befragten. Ein Mittelwert ist anders zu bewerten, wenn die zugrunde liegenden studentischen Urteile in großer Zahl übereinstimmen, als wenn sie sehr heterogen sind.

Die Antworten auf offene Fragen können graphisch zunächst nur schlecht dargestellt werden. Die einzelnen Aussagen können aber zusammengetragen und in zusammengehörende Themenblöcke gruppiert werden. So kann z.B. festgestellt werden, ob einzelne Kritikpunkte häufig wiederholt werden oder ob etwas als besonders gut oder schlecht empfunden wird.

Nun kommt es darauf an, aus den zusammengestellten Ergebnissen Konsequenzen zu ziehen. Es gilt dabei das Prinzip, die Stärken der Veranstaltungen beizubehalten und auszubauen und die Schwächen der Veranstaltungen soweit wie möglich auszuräumen. Bei einzelnen Aspekten kann das mit relativ einfachen Mitteln umgesetzt werden, beispielsweise wenn bemängelt wird, dass der Dozent zu leise oder zu schnell spricht oder wenn sich auffällig viele Teilnehmer mehr Informationen zu einem bestimmten Thema, wie z.B. zur Suche nach Lehrbüchern, wünschen. Der Dozent kann darauf reagieren, indem er in den Folgeveranstaltungen lauter bzw. langsamer spricht oder indem er ausführlicher auf die Recherche nach Lehrbüchern eingeht.

Wenn aber grundlegende Charakteristika der Veranstaltung, wie beispielsweise eine chaotische und verwirrende Struktur oder ein zu hoher Schwierigkeitsgrad der Übungsaufgaben auffällig häufig kritisiert werden, sollte durchaus das Gesamtkonzept der Veranstaltung überdacht und den Ansprüchen der Teilnehmer angepasst werden. Als besonders wertvoll und wichtig hat sich an der UB Würzburg herausgestellt, die Ergebnisse innerhalb der Arbeitsgruppe zu besprechen. Im Gespräch können Ursachen für ein schlechtes Evaluationsergebnis ermittelt und diskutiert werden. Außerdem ist der Austausch von Erfahrungen, die bei Schulungen gemacht wurden, hilfreich, um Kritikpunkte mit einfachen Mitteln abzuwenden.

Eine wichtige Grundhaltung in dieser Phase ist, bemängelte Punkte nicht vorrangig als „unveränderbare Defizite“ zu sehen, sondern als Chance zur Weiterentwicklung. Dies gilt für die strukturellen Aspekte der Veranstaltung, von denen bereits einige genannt wurden. Eine ganz besonders wichtige Rolle spielt diese Grundhaltung aber bei den Erkenntnissen, die sich aus der Beurteilung der lehrenden Person ziehen lassen, also bei den Einschätzungen der Dozentenvariablen. Persönliche Weiterentwicklung ist eine anspruchsvolle Aufgabe. Jeder Dozent hat Stärken und Schwächen, und es kann kein Evaluationsziel sein, am Ende ausschließlich den „perfekten Lehrenden“ kreieren zu wollen. Solch unrealistische Ziele haben Handlungstheorien zufolge häufig die Konsequenz, dass der Mensch gar nicht mehr agiert. Die von Sohr angesprochene Konsequenzlosigkeit wäre also gegeben. Realistischer ist es, dass jeder Lehrende mit Hilfe der Evaluationsergebnisse darauf hinarbeitet, seine persönlichen Stärken auszubauen und seine Schwächen bis zu einem gewissen Grad zu verändern. Persönliche Stärken und Schwächen sehen dabei aber sehr unterschiedlich aus. Für eine Veränderung scheint es günstig zu sein, wenn nicht ein Ideal von allen Bibliothekaren angestrebt werden soll.

Im Zusammenhang mit der Evaluation bzw. mit der Lehrtätigkeit an sich sollten für die Dozenten ausreichende Fortbildungsmöglichkeiten bestehen. Die Lehrenden werden sich einer Evaluation eher wirklich stellen, wenn sie ihre Kompetenzen selbst als gut einschätzen oder wenn sie meinen, die Möglichkeiten zu haben, sich zu verbessern. Fortbildungen sind sowohl im inhaltlichen Bereich z.B. bei neuen Veranstaltungen wichtig, ebenso aber bei didaktischen Fragen und persönlichen Fähigkeiten (Auftreten, Kommunikation ...).

Im Hinblick auf das Veränderungspotential spielt auch die Motivation für die Evaluation selbst eine Rolle: Sind evaluative Maßnahmen ein „Muss“, d. h. von der Direktion verordnet, dann ist „Konsequenzlosigkeit“ wiederum wahrscheinlich. Veränderungen sind eher dann zu erwarten, wenn die Evaluation vorrangig aus dem Willen nach Qualitätsverbesserung bei den betroffenen Mitarbeitern entsteht.

Letztlich halten wir die Evaluation bibliothekarischer Veranstaltungen für sinnvoll und notwendig. Eine gewissenhaft und kontinuierlich stattfindende Evaluation

ist mit großem Aufwand verbunden und bedeutet, sich offen mit Kritik auseinandersetzen zu müssen. Diesem Aufwand steht aber ein beträchtlicher Nutzen gegenüber: Evaluation hilft erstens, bibliothekarische Veranstaltungen zu verbessern. Zweitens ermöglichen die Ergebnisse den Bibliothekaren, das oft eher unbekannte Wesen „Benutzer“ besser kennen zu lernen. Für eine Serviceeinrichtung ist es generell wichtig, einen Eindruck von den Wünschen, Meinungen und Einstellungen der Nutzer zu haben. Und drittens hat Evaluation eine politische Dimension, sie dient auch der Eindruckssteuerung und Selbstdarstellung einer Bibliothek. Gerade für Bibliothekare ist die Dozentenrolle noch immer nicht selbstverständlich. Es kommt durchaus vor, dass Hochschulangehörige Zweifel an der Lehrkompetenz der Mitarbeiter anmelden, die dem „nichtwissenschaftlichen Personal“ zugerechnet werden. Eine methodisch sauber durchgeführte Lehrevaluation ist dabei eine stichhaltige Argumentationshilfe, die unabhängig von ihren Ergebnissen zeigt, dass Qualitätsstandards eingehalten werden.