

Birgit Hawelka, Stephanie Hiltmann:

Teaching Analysis Poll – ein Kodierleitfaden zur Analyse qualitativer Evaluationsdaten

Zusammenfassung

,Teaching Analysis Poll‘ ist als qualitative Methode der formativen Lehrveranstaltungsevaluation inzwischen weit verbreitet. Die Stärken dieser moderierten Gruppendiskussion liegen vor allem in der strukturoffenen Fokussierung des Lernprozesses der Studierenden sowie der Verknüpfung der Lehrveranstaltungsevaluation und Lehrberatung.

Aus methodischer Sicht problematisch sind bei ,Teaching Analysis Poll‘ bislang die Analyse und Interpretation der erhobenen Daten. Die Studierenden äußern subjektive Sichtweisen in Bezug auf die jeweilige Lehrveranstaltung, die dann als Grundlage für ein Beratungsgespräch unter didaktischen Gesichtspunkten analysiert und interpretiert werden müssen. Als Beitrag zur Standardisierung dieser Prozesse wurde an der Universität Regensburg ein Kodierleitfaden entwickelt.

In diesem Artikel wird die Entwicklung des Leitfadens beschrieben und begründet. Zugleich wird skizziert, wie sich die Durchführung von ,Teaching Analysis Poll‘ auf der Grundlage des Leitfadens verändert.

Gliederung

- 1. ,Teaching Analysis Poll‘ als Methode der qualitativen Lehrveranstaltungsevaluation**
- 2. Diskussion der Vorgehensweise**
- 3. Kategoriensystem zur Bewertung von Lehre**
 - 3.1 Lehre als Unterstützung von Lernprozessen*
 - 3.2 Bewertung von Lehr-Lern-Prozessen*
 - 3.3 Ableitung weiterer Kategorien aus Evaluationsbogen*
- 4. Entwicklung eines Kodierleitfadens**
- 5. ,Teaching Analysis Poll‘ auf der Basis des Kodierleitfadens**
- 6. Zusammenfassung und Diskussion**
- 7. Literatur**

1. ,Teaching Analysis Poll‘ als Methode der qualitativen Lehrveranstaltungsevaluation

Lehrveranstaltungsevaluationen sind an Hochschulen inzwischen fester Bestandteil der systematischen Qualitätssicherung von Lehre. Sie dienen der Wirkungskontrolle oder verfolgen das Ziel, Lehr- und Lernprozesse zu reflektieren und konstruktive Anhaltspunkte zur Verbesserung von Lehrveranstaltungen zu gewinnen.

Gerade wenn studentische Rückmeldungen zur Fortentwicklung von Lehrveranstaltungen genutzt werden sollen, ist ,Teaching Analysis Poll‘ als Methode zur qualitativen Zwischenevaluation gut geeignet. Im Gegensatz zu vielen anderen qualitativen Ansätzen erfordert sie nur einen geringen Vorbereitungsaufwand und läuft nach einem einfachen Schema ab. Die Lehrperson beendet ihre Lehrveranstaltung zwanzig Minuten früher und verlässt den Raum. Ein externer Moderator führt nun die Befragung durch (Frank, Fröhlich & Lahm, 2011).

Da ,Teaching Analysis Poll‘ in Form einer moderierten Gruppendiskussion stattfindet, teilen sich die Studierenden in Kleingruppen auf. Der Moderator stellt den Studierenden lediglich drei Fragen:

1. Was unterstützt in dieser Veranstaltung Ihren Lernprozess?
2. Welche Aspekte erleben Sie als eher hinderlich?
3. Wodurch könnte Ihrer Meinung nach Ihr Lernprozess noch stärker unterstützt werden?

Die Studierenden erhalten zehn Minuten Zeit, um diese Fragen in ihren Kleingruppen zu diskutieren.

Nach der Diskussion halten die Studierenden ihre Ergebnisse in Stichpunkten schriftlich fest.

Anschließend werden im Plenum die genannten Punkte der Kleingruppen nochmals von der Moderation paraphrasiert und wiedergespiegelt.

Ein erstes Feedback zur Einschätzung der Studierenden erhalten die Lehrpersonen möglichst zeitnah nach der Erhebung. In einem anschließenden Beratungsgespräch werden die Evaluationsdaten als Grundlage für eine Lehrberatung herangezogen.

2. Diskussion der Vorgehensweise

Die skizzierte Methodik von ,Teaching Analysis Poll‘ weist im Vergleich zu anderen Evaluationsansätzen zwei wesentliche Vorteile auf:

,Teaching Analysis Poll‘ erfasst Qualitätsmerkmale kontextbezogen. Lehrveranstaltungen sind in Teilnehmerzusammensetzung, Durchführung und Zielsetzung sehr heterogen. Entsprechend kann die Festlegung und Operationalisierung von Evaluationskriterien (als Determinanten erfolgreichen didaktischen Handelns) nicht allgemeingültig, sondern nur vor dem Hintergrund des spezifischen Kontexts erfolgen (Renkl, 2015). Als strukturoffene Befragungsmethode wird ,Teaching Analysis Poll‘ dieser Anforderung in stärkerem Maße gerecht als viele standardisierte Evaluationsverfahren (Hawelka & Wild, 2014).

Evaluation und Lehrberatung sind miteinander verknüpft. Es gibt einiges an empirischer Evidenz dafür, dass Evaluation per se noch zu keiner Steigerung der Lehrqualität führt (Rindermann, 2003). Nennenswerte Effekte sind nur dann zu verzeichnen, wenn die Evaluation mit einer Intervention verbunden wird. Penny und Coe (2004) berechneten in einer Metaanalyse eine Effektstärke von $d+ = 0.69$ bezüglich der Verbesserung der Lehre, für den Fall, dass studentisches Feedback mit einer Lehrberatung verbunden wird. Bei ‚Teaching Analysis Poll‘ wird der Lernprozess in einer Lehrveranstaltung zunächst aus studentischer Perspektive analysiert und im Rahmen der Methode systematisch mit einem Beratungsgespräch verknüpft. Die Evaluation bildet somit die Grundlage für eine datenbasierte Lehrberatung.

Gleichzeitig sind an der eingangs vorgestellten Vorgehensweise einige Punkte als kritisch zu betrachten:

Fehlende Berücksichtigung der Lernziele einer Veranstaltung. Die Bewertung einer Lehrveranstaltung und damit einhergehend eine Entwicklung von Verbesserungsansätzen kann nur vor dem Hintergrund der angestrebten Lernziele erfolgen. Die Auswahl des Formats, der Methoden, der Lernmaterialien usw. muss von den gesetzten Lernzielen einer Veranstaltung bestimmt werden. Nicht die Anzahl und hohe Variabilität an Lehrmethoden zeichnen eine gute Lehrveranstaltung aus, sondern der Umstand, inwieweit die eingesetzten Methoden für die Erreichung der Lernziele geeignet sind und wie gut sie umgesetzt werden (Renkl, 2015). Für eine fundierte Interpretation der aus ‚Teaching Analysis Poll‘ gewonnenen Daten ist es daher unerlässlich, die intendierten Lernziele einer evaluierten Lehrveranstaltung zu kennen.

Didaktische Analyse während der Befragung. Das oben skizzierte prototypische Vorgehen bei ‚Teaching Analysis Poll‘ beinhaltet die Frage nach Verbesserungsvorschlägen. Dabei nennen Studierende Strategien zur Veränderung der Lehrveranstaltung und (vermeintlichen) Verbesserungen als hinderlich erlebten Punkte. Solche Ad-hoc-Lösungen, die innerhalb weniger Minuten während der Datenerhebung entwickelt werden, bergen allerdings die Gefahr von subjektiv geprägten Fehlinterpretationen und verkürzten Lösungsansätzen. So kann aus studentischer Perspektive beispielsweise das Problem ‚Texte sind zu schwierig‘ zum vordergründig naheliegenden Lösungsvorschlag ‚Einfachere Texte bereitstellen‘ führen. Aus didaktischer Perspektive hilfreicher wäre in diesem Beispiel aber möglicherweise die Ergänzung der Texte mit lernleitenden Fragen, die den Lese- und Lernprozess organisieren und fokussieren. Fundierte Ansätze zur Fortentwicklung von Lehrveranstaltungen können in der Regel erst nach sorgfältiger und strukturierter Analyse aller Schwachpunkte einer Lehrveranstaltung und vor dem Hintergrund der Lernziele gezogen werden.

Die Qualität der Analyse ist abhängig vom Moderator. Studentische Rückmeldungen zur Lehrqualität geben in erster Linie Einschätzungen und subjektive Wahrnehmungen im Sinne von

„Alltags-Evaluationen“ (Kromrey, 2001, S. 20) wieder. Sie können damit wertvolle Anregungen zur Reflexion geben, der Lehrperson einen Perspektivenwechsel ermöglichen sowie ‚blinde Flecken‘ aufdecken. Fraglich ist jedoch, wie (konstrukt-)valide die Nennungen der Studierenden sind. Als Grundlage für ein Beratungsgespräch müssen daher die studentischen Rückmeldungen unter didaktischen Gesichtspunkten strukturiert, analysiert und interpretiert werden. Andernfalls erhalten die Lehrenden lediglich eine Beschreibung ihrer Lehrveranstaltung aus studentischer Sicht, nicht aber eine Bewertung der Lehrveranstaltung nach didaktischen Maßstäben. Welche Maßstäbe und Kriterien dabei herangezogen werden, obliegt in der Praxis bislang allein der Person des Evaluators. Ein Rückbezug der erhobenen Daten auf empirisch belegte Evidenzen guter Lehre bzw. lernpsychologisch fundierte Theorien kann zwar vermutlich oftmals unterstellt, nicht aber gewährleistet werden. Gerade diese Rückkopplung bei der Analyse der Daten ist aber nötig, wenn die Methode beansprucht, ein Schnittpunkt zwischen einer Evaluation, die sozialwissenschaftlichen Standards genügt, und einer datenbasierten Lehrberatung zu sein. „Um eine zuverlässige Methode zu entwickeln, die eine vermeintliche Strukturlosigkeit des Gesagten vermeidet, ist der Diskurs [in Gruppendiskussionen] nach bestimmten Kriterien ... zu analysieren“ (Ernst, 2008, S. 63). Die Entwicklung und Begründung solch eines Kriterienkatalogs zur Strukturierung von Rückmeldungen ist Gegenstand des folgenden Kapitels (3.).

3. Kategoriensystem zur Bewertung von Lehre

Kern eines Kodierleitfadens ist ein System zur Kategorisierung und Strukturierung der Rückmeldung. Solch ein Kodierungssystem für ‚Teaching Analysis Poll‘ muss – schon allein aufgrund der Fragestellung bei dieser Erhebungsmethode – die Unterstützung des Lernprozesses von Studierenden in besonderem Maße berücksichtigen.

3.1 Lehre als Unterstützung von Lernprozessen

Voraussetzung für die Entwicklung von Bewertungskriterien ist eine Vorstellung davon, was unter guter Lehre verstanden werden kann. Dabei kann es nicht darum gehen, objektive und kontextunabhängige Qualitätsmerkmale zu definieren. Dimensionen der Qualität erschließen sich vielmehr in der Auseinandersetzung von Studierenden mit einem bestimmten Lerninhalt (Kromrey, 2001). Entsprechend misst sich Lehrqualität in erster Linie daran, wie gut es Lehrenden gelingt, geeignete Lernprozesse von Studierenden zu initiieren und zu unterstützen, um ein gesetztes Lernziel zu erreichen.

Ausgangspunkt des Lernprozesses ist die *Bereitstellung und Auswahl von Lernaufgaben*. Dies ist die erste und möglicherweise wichtigste „Schnittstelle zwischen Dozent, Student und Curriculum“

(Wild, 2000, S. 121). Wesentlich für die Qualitätsbeurteilung des Lernprozesses ist, dass die Lernaufgaben geeignet sind, das Lernziel zu erreichen.

Den Prozess, der dann in Auseinandersetzung mit einer Lernaufgabe zur Anwendung von Lernstrategien führt, beschreibt das Dreiphasenmodell der Lernstrategienutzung im Studium („3PLS-Modell“) von Wild (2000):

Demnach erfolgt nach der ersten Auseinandersetzung mit der Lernaufgabe ihre *subjektive Konstruktion*. Zu Beginn des Lernprozesses bilden sich die Studierenden eine Vorstellung davon, welche Lernleistung von ihnen erwartet wird und welche kognitiven Strategien geeignet sind, um die Lernziele zu erreichen. Lehrende können den Studierenden auf unterschiedliche Art und Weise signalisieren, welche Erwartungen sie mit der Bearbeitung einer Lernaufgabe seitens Studierender verbinden. Als zwei wichtige Punkte nennt hier Wild (2000) Handlungsanweisungen und die Modellfunktion von Lehrenden und Kommilitonen bei der Aufgabenbearbeitung. Eine weitere zentrale Funktion dürfte hier die explizite Nennung der Lernziele einnehmen, die sich als handlungsleitend für Lernprozesse erwiesen hat (Biggs & Tang, 2011).

In einem zweiten Schritt, der *motivationalen Regulation*, entscheiden die Studierenden, ob sie sich mit der Aufgabe auseinandersetzen und welche Lernstrategie sie anwenden wollen. Die Bereitschaft dazu hängt von persönlichen Zielen und Werten sowie der subjektiven Einschätzung der Erfolgswahrscheinlichkeit der geforderten Lernhandlung ab. – Wie sich Studierende dann tatsächlich mit einem Lerninhalt auseinandersetzen und welche Lernstrategien sie anwenden, entscheidet sich in der Phase der *(Lern-)Handlungsregulation*. Die Einflussfaktoren der hier getroffenen Entscheidungen sind das Gegenstück zu denen in der Phase der motivationalen Regulation. Abhängig vom tatsächlichen Vorwissen und von den kognitiven Fähigkeiten (Personenmerkmale) oder tatsächlich vorhandenen Unterstützungsleistungen (Kontextfaktoren) werden Lernstrategien zur Bearbeitung der subjektiv konstruierten Aufgabe ausgewählt und eingesetzt (Wild, 2000). Der Unterschied in beiden Phasen besteht im Wesentlichen darin, dass die Motivationsregulation durch antizipierte, die Handlungsregulation durch tatsächlich erlebte Person- und Kontextfaktoren geprägt sind. Müller (2007) fasst auf der Basis verschiedener lern- und motivationstheoretischer Zugänge die Einflussmöglichkeiten Lehrender auf die Motivation von Studierenden zusammen und hebt dabei sechs wesentliche Komponenten hervor: Autonomie- und Kompetenzunterstützung, soziale Einbindung, sowie Interesse der Lehrenden, Relevanz der Inhalte und Instruktionsqualität.

Die kognitive Verarbeitung der Lernaufgabe schließlich erfolgt durch den Einsatz *kognitiver Lernstrategien*. Wild und Schiefele (1994) unterscheiden 4 Strategien: a) Organisationsstrategien, um den Lerninhalt neu zu organisieren und zu veranschaulichen, b) Wiederholungsstrategien,

welche die Wissensaufnahme unterstützen und auf das Einprägen des Lerninhalts abzielen, c) Elaborationsstrategien zur Einbettung neuen Wissens in bereits bestehende Vorwissensstrukturen und d) Strategien des kritischen Prüfens und Hinterfragens des Lerninhalts. Abhängig von der jeweiligen Aufgabe, der gegebenen Lernmotivation und den kognitiven Fähigkeiten werden verschiedene Lernstrategien unterschiedlich intensiv angewendet. Lehrende können die Anwendung von Lernstrategien aber unterstützen, indem sie Studierende in Lehrveranstaltungen durch entsprechende Aufgabenstellungen explizit dazu anregen, beispielsweise Lerninhalte zu wiederholen, zu elaborieren oder kritisch zu prüfen.

Der alleinige Einsatz kognitiver Strategien erreicht aber dann seine Grenzen, wenn es – wie im Studium – um die Erreichung höherer Lernziele geht (Hasselhorn & Labuhn, 2017). Mit zunehmender Komplexität der Ziele steigt die Notwendigkeit, dass die Lernaktivitäten geplant, hinsichtlich ihrer Wirksamkeit kontrolliert und ggf. modifiziert werden. In klassischen didaktischen Modellen, (vorwiegend aus dem schulischen Bereich; z. B. Meyer, 2016) wird diese Aufgabe der *Steuerung des Lernprozesses* den Lehrpersonen zugeschrieben. Abhängig von den Merkmalen der Studierenden oder auch dem Schwierigkeitsgrad der Lernaufgabe ist es für den Hochschulbereich sinnvoll und sogar wünschenswert, dass die Steuerung des eigenen Lernprozesses außerhalb der Lehrveranstaltung durch den Einsatz metakognitiver Strategien sukzessive durch die Lernenden selbst übernommen wird.

Damit Lernen erfolgreich sein kann, ist zudem die Aktivierung von *Ressourcen* notwendig, die den Lernprozess unterstützen und vor äußeren Einflüssen abschirmen. Dazu zählen einerseits *innere Ressourcen*, z. B. die eigene Anstrengungsbereitschaft oder die investierte Lernzeit, andererseits auch *äußere Ressourcen*, wie die Gestaltung der Lernumgebung oder die Nutzung zusätzlicher Informationsquellen (Wild & Schiefele, 1994).

Erst das adäquate Zusammenspiel von kognitiven, ressourcenbezogenen und metakognitiven Lernstrategien (bzw. der externen Steuerung von Lernprozessen) ermöglicht einen erfolgreichen Lernprozess.

Aufgrund der dargestellten Überlegungen zur Unterstützung des Lernprozesses wurden als *Kategorien* die einzelnen Phasen vom Bereitstellen einer Lernaufgabe bis hin zum Erreichen des Lernziels festgelegt (siehe Abbildung 1). Da die Phasen der Motivations- und Handlungsregulation durch dieselben Kontextfaktoren unterstützt werden, wurden diese beiden Phasen zu einer Kategorie, ‚Motivation‘, zusammengefasst. So entstand ein Gerüst von 7 Kategorien.

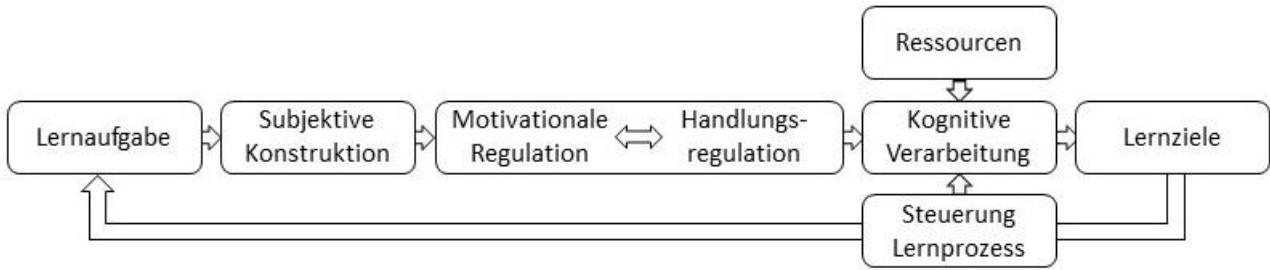


Abbildung 1. Kategorien im Bereich ‚Lernen‘ (eigene Darstellung, in Anlehnung an das 3PLS-Modell, von Wild, 2000)

3.2 Bewertung von Lehr-Lern-Prozessen

Die Ausführungen zur Unterstützung von Lernprozessen könnten vermutlich je nach theoretischem Zugang beliebig ergänzt werden. Für die Evaluation von Lehr-Lern-Prozessen wesentlich ist aber, dass die jeweiligen Unterstützungsmöglichkeiten auch tatsächlich von den Studierenden und Lehrenden als solche erkannt und damit als Bewertungskriterien akzeptiert werden. Aspekte guter Lehre werden in validen Evaluationsbogen erhoben. Für die weitere Operationalisierung der beschriebenen Kategorien (siehe Abschnitt 3.1) wurde daher auf bewährte Evaluationsbogen zurückgegriffen.

Obwohl verschiedene Instrumentarien dasselbe Ziel – nämlich die Messung von Lehrqualität – verfolgen, ist festzustellen, dass diese, abhängig vom zugrunde gelegten Forschungszugang, inhaltlich divergieren (Wolbring, 2013). Um ein möglichst breites Spektrum des Konstruktus guter Lehre abzudecken, wurden daher als Grundlage für die Entwicklung eines Kategoriensystems drei in ihrer Entwicklung und Herangehensweise sehr verschiedene, standardisierte Evaluationsinstrumentarien herangezogen:

1. ‚Students‘ Evaluations of Educational Quality (SEEQ) von Marsh (1982), der als Quasi-Standard im amerikanischen Hochschulraum gelten kann,
2. ‚Heidelberger Inventar zur Lehrveranstaltungsevaluation in verschiedenen Entwicklungsstufen‘ (HILVE) von Rindermann (2009), der im deutschsprachigen Raum weit verbreitet ist, und
3. ‚Fragebogen zum Dozierendenverhalten‘ (FRADOV) von Koch (2004), der sich für diesen Zweck deshalb besonders eignet, weil er weniger empirisch geprägt, sondern vielmehr konsequent aus instruktions- und kognitionspsychologischen Theorien abgeleitet ist.

Als Ausgangsmaterial für die weitere Analyse wurden die insgesamt 50 Skalen der drei Evaluationsinstrumentarien verwendet. Diese Skalen wurden inhaltlich nach ihrer Funktion zur Unterstützung des Lernprozesses analysiert und dann deduktiv den 7 Kategorien des Bereichs ‚Lernen‘ zugeordnet. Insgesamt konnten so 31 von 50 Skalen eingeordnet werden.

Für die Kategorien, die durch mehrere, inhaltlich heterogene Skalen besetzt waren, wurden in einem weiteren Schritt aus den Skalen Subkategorien gebildet (siehe Tabelle 1).

Tabelle 1

Entwicklung eines Kategoriensystems für den Bereich ‚Lernen‘

Kategorien zum Bereich ‚Lernen‘	Deduktive Zuordnung der Skalen aus den Evaluationsbögen	Induktive Ableitung von Subkategorien
Lernaufgabe	–	–
Subjektive	–	–
Konstruktion der Lernaufgabe		
Motivationale	Autonomie (F)	Autonomie
Regulation	Flexibilität (F)	
	Kompetenzunterstützung (F)	Kompetenzerleben
	Anforderungen (H)	ermöglichen
	Verhältnis Dozierende/Studierende (F)	Soziale Eingebundenheit
	Klima (H)	
	Beziehung zum Einzelnen (S)	
	Enthusiasmus (F)	Interesse und Engagement der
	Dozentenengagement (H)	Lehrenden
	Enthusiasmus (S)	
	Didaktisch-konstruktivistische Lehraspekte (F)	Interessantheit und Relevanz
	Interessantheit (H)	
Kognitive	Wiederholung (F)	Wiederholung
Verarbeitung	Organisation (F)	Lerninhalte organisieren
	Aufmerksamkeit wecken (F)	Elaboration

	Elaboration (F)	
	Auseinandersetzung (H)	
	Breite (S)	
	Stimulierung (F)	Inhalte kritisch prüfen
	Kritisches Prüfen (F)	
	Verarbeitung (H)	
Steuerung des Lernprozesses	Planung (F)	Lehre planen und strukturieren
	Struktur (H)	
	Überprüfung (F)	Lernfortschritt überprüfen
	Regulierung (F)	Lehrprozess adaptiv regulieren
Ressourcen	Betreuung (H)	Ansprechpartner bei Problemen
	Aufgaben (S)	Lernmaterialien/Literatur
Lernerfolg	Lernen–quantitativ (H)	–
	Lernen–qualitativ (H)	
	Interessenförderung (H)	
	Lernen (S)	

Anmerkungen. F = FRADOV; H = HILVE; S = SEEQ

3.3 Ableitung weiterer Kategorien aus Evaluationsbogen

Aus 17 der 19 verbliebenen Skalenⁱ der Fragebogen wurden induktiv 4 weitere Kategorien gebildet: „Interaktion“, „Gesamtbewertung“, „Personenmerkmale“ und „Prüfen“.

Aufgrund der hohen Anzahl an Skalen in der Kategorie „Interaktion“ wurde dieser Bereich weiter differenziert. Eine inhaltliche Analyse der Skalen dieses Bereichs ergab 3 Subkategorien:

- a) „Verständliche und anregende Darstellung der Inhalte“, b) „Aktive Beteiligung und Interaktionsmanagement“ und c) „Veranstaltung leiten“ (siehe Tabelle 2).

Tabelle 2

Ableitung weiterer Kategorien anhand der Skalen

Skalen aus den Evaluationsbogen	Ableitung von	Ableitung von Subkategorien
	Kategorien	
Variabilität (F)	Interaktion	Verständliche und anregende
Allgemeine Verständlichkeit (F)		Darstellung der Inhalte
Lehrkompetenz (H)		
Organisation (S)		
Kommunikative Unterrichtsformen (H)		Aktive Beteiligung und
Interaktionsmanagement (H)		Interaktionsmanagement
Diskussion (F)		
Beteiligung (H)		
Interaktion in der Gruppe (S)		
Anomie (H)		Veranstaltung leiten
Lenkung (F)		
Gesamtbewertung (S)	Gesamtbewertung	–
Allgemeineinschätzung (H)		
Redundanz (H)	Personenmerkmale	–
Fleiß/Arbeitshaltung (H)		
Thema (H)		
Prüfungen (S)	Prüfen	

Anmerkungen. F = FRADOV; H = HILVE; S = SEEQ

Insbesondere bei der Kategorie ‚*Interaktion*‘ deutete sich an, dass die Unterscheidung der einzelnen Kategorien nicht immer trennscharf gelingt. Einige Skalen der Kategorie ‚*Interaktion*‘ erfassen einerseits sichtbare Interaktionsformen (z. B. die Gelegenheit zur Diskussion), implizieren aber gleichzeitig die Unterstützung einer Lernstrategie (z. B. das kritische Prüfen eines Gegenstands durch Diskussionen). Dass diese Problematik der fehlenden Trennschärfe insbesondere den Bereich

Interaktion betrifft, ist wenig verwunderlich, da Interaktion in Lehrveranstaltungen kein Selbstzweck ist, sondern immer ein didaktisches Ziel verfolgt.

4. Entwicklung eines Kodierleitfadens

Kern des Kodierleitfadens bildet das Kategoriensystem. Zur Überprüfung der Tauglichkeit des Systems zur Analyse von Erhebungen aus ‚Teaching Analysis Poll‘ wurden zunächst die Ergebnisse von 58 Erhebungen aus verschiedenen Fakultäten und Veranstaltungsformaten der Universität Regensburg in das Kategorienschema eingeordnet, wobei eine studentische Aussage dabei eine Kodiereinheit bildete.

Es zeigte sich einerseits, dass die 3 Kategorien ‚Lernaufgabe‘, ‚Prüfungen‘ und ‚Personenmerkmale‘ durch Rückmeldungen nicht bzw. kaum besetzt waren. Diese Kategorien wurden daher aus dem Katalog gestrichen. Andererseits machten die Studierenden häufig Aussagen zu Rahmenbedingungen von Lehrveranstaltungen (wie z. B. dem Raum, der Zeit etc.). Diese Kategorie wurde daher – zusätzlich zu den theoretisch bzw. empirisch belegten Kriterien – mit aufgenommen.

Ergebnis dieses Revisionsprozesses ist ein Kategoriensystem mit 9 Hauptkategorien und 17 Subkategorien (siehe Tabelle 3).

Tabelle 3

Überblick über die Kategorien

Kategorie	Subkategorie
1 Interaktion	1.1 Verständliche und anregende Darstellung der Inhalte 1.2 Aktive Beteiligung und Interaktionsmanagement 1.3 Veranstaltung leiten
2 Subjektive Konstruktion der Lernaufgabe	
3 Motivationale Regulation	3.1 Autonomie 3.2 Kompetenzerleben ermöglichen 3.3 Soziale Eingebundenheit 3.4 Interesse und Engagement der Lehrenden

3.5 Interessantheit und Relevanz

4 Kognitive Verarbeitung

4.1 Wiederholen

4.2 Lerninhalte organisieren

4.3 Elaboration

4.4 Inhalte kritisch prüfen

5 Steuerung des Lernprozesses

5.1 Lehre planen und strukturieren

5.2 Lernfortschritt überprüfen

5.3 Lehrprozess adaptiv regulieren

6 Ressourcen

6.1 Ansprechpartner bei Problemen

6.2 Lernmaterialien/Literatur

7 Lernerfolg

—

8 Gesamтурteil

—

9 Rahmenbedingungen

—

Jede Kategorie kann dabei nur zwei Ausprägungen erhalten: Entweder das angesprochene Thema wirkt unterstützend oder eher hinderlich.

Auf den Kategorien aufbauend wurde ein Kodierleitfaden erstellt, der für alle Kategorien Definitionen und Kodierregeln enthält. Gleichzeitig wurden anhand des Datenmaterials Ankerbeispiele formuliert. Abbildung 2 zeigt dies exemplarisch für die Subkategorie 3.1 „Autonomie“.

3 Motivationale Regulation

3.1 Autonomie

Die Studierenden erhalten Freiräume, um ihren Lernprozess innerhalb und außerhalb der Lehrveranstaltung autonom zu gestalten.

Kodiert werden in dieser Kategorie

- Studierende können eigene Interessen und Ideen in den Lernprozess einbringen
- Studierende erhalten die Möglichkeit, selbständig nach Lösungsmöglichkeiten für Probleme zu suchen
- Studierende können Lehrmittel, Unterrichts- oder Sozialformen mitbestimmen

Ankerbeispiele

für hilfreiches Verhalten	für hinderliches Verhalten
<ul style="list-style-type: none">(+) selbständiges Arbeiten(+) eigenverantwortliche Gestaltung der Stunden(+) Freiheit bei der Themenwahl(+) Dozent ist kompromissfähig(+) Wahlfreiheit bei Schwerpunkt des Referats(+) viel Freiraum (gut bei Projekten)	<ul style="list-style-type: none">(-) zu wenig Freiraum bei der Aufgabenbearbeitung(-) Dozent ist ungeduldig, greift bei Fehlern sofort ein und bringt Lösung

Nicht kodiert werden in dieser Kategorie

- Lehrperson reagiert flexibel bei Verständnisschwierigkeiten \Rightarrow (5.3) Lehrprozess adaptiv regulieren
- Anwesenheitspflicht \Rightarrow (9) Rahmenbedingungen

Abbildung 2. Exemplarische Darstellung des Kodierleitfadens für die Subkategorie 3.1

,Autonomie‘ (Hawelka, 2017, S. 9)

In einer Qualitätskontrolle mit 75 Rückmeldungen und 1.774 Kodiereinheiten erwies sich der Kodierleitfaden als gut handhabbar. 97,4 % aller Aussagen konnten den Kategorien zugeordnet werden. Der Leitfaden trug zur klaren Kodierung bei, was sich auch in einer hohen Intercoder-Reliabilität (Cohens $\kappa = .82$) widerspiegelt. Aufgrund der engen Orientierung an validierten Instrumentarien kann eine Konstruktvalidität des Leitfadens angenommen werden.

5. ,Teaching Analysis Poll‘ auf der Basis des Kodierleitfadens

Unter Berücksichtigung der skizzierten kritischen Punkte beim typischen Ablauf von ,Teaching Analysis Poll‘ (siehe Kapitel 2.) und des entwickelten Kodierleitfadens (siehe Kapitel 4.) ergibt sich ein modifizierter Ablauf (siehe Abbildung 3).

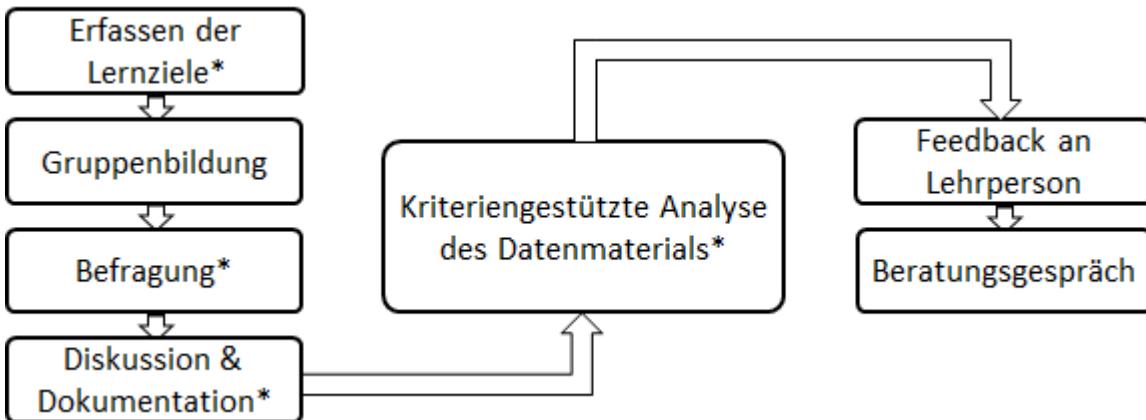


Abbildung 3. Modifizierter Ablauf von ‚Teaching Analysis Poll‘

Die mit * gekennzeichneten Schritte sind im Vergleich zum klassischen Vorgehen neu bzw. modifiziert.

1. *Erfassen der Lernziele*: Da die Bewertung von Evaluationsergebnissen nur vor dem Hintergrund der intendierten Lernziele erfolgen kann, ist es notwendig, die Lernziele der zu evaluierenden Lehrveranstaltung zu kennen. Sinnvoll scheint es, bereits vor der Erhebung (z. B. bei der Terminfindung) die von der Lehrperson gesetzten Lernziele abzufragen.

2. *Gruppenbildung*: ‚Teaching Analysis Poll‘ wird in Form einer moderierten Gruppendiskussion durchgeführt. Die Studierenden teilen sich dazu auf Kleingruppen von drei bis sieben Personen auf. Eine Anzahl von drei bis fünf Gruppen reicht dabei aus, um eine *theoretische Sättigung* zu erreichen. Eine höhere Anzahl an Gruppen wird in der Regel keinen Mehrwert an Erkenntnissen bringen (Morgan, 2009). Daher kann in größeren Lehrveranstaltungen mit mehr als 35 Teilnehmenden eine zufällige Stichprobe an Studierenden zur Erhebung ausgewählt werden.

3. *Befragung*: Zur Evaluation der Lehrveranstaltung werden den Studierenden die Evaluationsfragen gestellt:

- Was unterstützt in dieser Veranstaltung Ihren Lernprozess?
- Welche Aspekte erleben Sie als eher hinderlich?

Aufgrund der skizzierten Problematik bei der Abfrage von Verbesserungsvorschlägen (siehe Kapitel 2.) wird an der Universität Regensburg auf die übliche dritte Frage verzichtet. Die Erhebung beschränkt sich auf die Analyse der Stärken und Schwächen und mögliche Konsequenzen daraus werden erst nach einer ausführlichen, kriteriengestützten Analyse (siehe Abbildung 3) abgeleitet.

4. *Diskussion und Dokumentation*: Die Studierenden diskutieren diese Fragen kurz in ihren Kleingruppen und halten die Ergebnisse schriftlich fest. Diese werden dann nochmals im Plenum besprochen. Die Moderation hat in diesem Schritt zwei Aufgaben, die bereits zur systematischen Analyse der studentischen Aussagen beitragen:

a) Interpretation der Rückmeldungen: Mehrdeutige und missverständliche Aussagen werden durch gezieltes Nachfragen vor dem Hintergrund der oben dargestellten Kategorien geklärt und die Rückantworten der Studierenden so in eine didaktische Sprache übersetzt. Beispielhaft kann dieses Vorgehen durch die Aussage ‚Nur Referate‘ erläutert werden. Diese Aussage lässt zunächst unterschiedliche Deutungsmuster zu: Einerseits könnten die Studierenden den Ablauf der Sitzungen als monoton empfinden, weil als didaktische Methode ausschließlich studentische Referate eingesetzt werden. Andererseits ist es denkbar, dass die Studierenden Input von der Lehrperson vermissen und den Input von Kommilitonen als zu umfangreich bewerten. In diesem Beispiel wurde durch Nachfragen der Moderation deutlich, dass die Studierenden es als hinderlich für ihren Lernprozess empfinden, dass die Lehrperson zu wenig Input gibt. Das Problem liegt in diesem Beispiel also in der fehlenden Modellierung der Lernaufgabe seitens der Lehrperson und nicht in einer mangelnden Interessantheit der Lehrveranstaltung.

Die Ergebnisse der *formulierenden Interpretation* (Loos & Schäffer, 2001) werden durch die Moderation direkt dokumentiert (siehe Abbildung 4).

Nur Referate = wenig Input Dozent

Abbildung 4. Beispiel für die Klärung mehrdeutiger Aussagen

b) Festlegung der Kodiereinheiten: In der Regel entspricht ein genannter Stichpunkt auch einer Kodiereinheit (z. B. ‚Nur Referate‘). Häufig nennen die Studierenden jedoch zwei oder mehrere inhaltliche Aspekte in einem Stichpunkt. Aufgabe der Moderation ist es daher, durch gezielte Rückfragen zu klären, ob sich die unterschiedlichen Aussagen auf ein und denselben Punkt beziehen oder ob damit unterschiedliche Aspekte der Lehre gemeint sind. Im unten abgebildeten Beispiel zur Aussage ‚Dozentin ist freundlich und hilfsbereit‘ ergab erst die Nachfrage der Moderation, dass damit zwei aus didaktischer Perspektive unterschiedliche Aspekte angesprochen werden: Zum einen tritt die Dozentin in der Lehrveranstaltung freundlich auf (Subkategorie 3.3 ‚Soziale Eingebundenheit‘). Gleichzeitig ist die Dozentin hilfsbereit bei der Vorbereitung von Referaten (Subkategorie 6.1 ‚Ansprechpartner bei Problemen‘). Folglich wurde seitens der Moderation durch eine Trennlinie kenntlich gemacht, dass es sich dabei um verschiedene Kodiereinheiten handelt (siehe Abbildung 5).

**Dozentin ist freundlich und hilfsbereit
Vorbereitung Referate**

Abbildung 5. Beispiel für die Festlegung der Kodiereinheiten

5. Kriteriengestützte Analyse des Datenmaterials. Die eigentliche Analyse der erhobenen Daten erfolgt dann außerhalb der Lehrveranstaltung. Mithilfe des vorliegenden Kodierleitfadens werden die studentischen Aussagen in ein Kategoriensystem übertragen und so nach didaktischen Kriterien sortiert. Der Leitfaden bildet somit eine einheitliche und intersubjektiv überprüfbare Grundlage der Analyse. Rückmeldungen, die zur Bewertung von Lernprozessen nicht valide sind, werden mit dem Kategoriensystem ausgesiebt. Die Bewertung der einzelnen Aussagen erfolgt im Anschluss an die Analyse und vor dem Hintergrund der erhobenen Lernziele. Studierende können eine Beobachtung wie beispielsweise ‚Häufiges Wiederholen‘ als negativen Aspekt der Veranstaltung eingeordnet haben, während eine Person mit didaktischer Expertise diese Rückmeldung im Hinblick auf ein mögliches Lernziel, wie ‚Studierende können die theoretischen Grundlagen detailliert wiedergeben‘, jedoch als positiv und wichtig bewertet. Eine Verbesserungsmöglichkeit könnte dann der Einsatz variierender Methoden zur Wiederholung der Inhalte sein.

6. Feedback an Lehrperson: Ein erstes Feedback über die Einschätzung der Studierenden erhalten die Lehrpersonen per E-Mail möglichst zeitnah nach der Erhebung. Durch eine Sortierung, die sich am Kodierleitfaden orientiert, können die Aussagen der Studierenden in strukturierter Form den jeweiligen Lehrpersonen rückgemeldet werden.

7. Beratungsgespräch: Das anschließende Feedbackgespräch wird in der Regel nach dem *beratenden Ansatz* durchgeführt (Penny & Coe, 2004). Dabei werden die Evaluationsdaten als Grundlage für das Gespräch herangezogen. Im Gespräch werden dann die Selbsteinschätzung der Lehrperson aufgenommen und besondere Gegebenheiten des jeweiligen Kurses diskutiert. Darauf aufbauend entwickeln der didaktische Berater und die Lehrperson gemeinsam Verbesserungsstrategien.

6. Zusammenfassung und Diskussion

Die Diskussion um die Evaluation mit ‚Teaching Analysis Poll‘ zeigte, dass diese Methodik der qualitativen Evaluation gut geeignet ist, um zur Verbesserung der Lehrqualität beizutragen. Wesentliches Qualitätsmerkmal bei der Anwendung der Methodik ist jedoch, dass die Rückmeldungen der Studierenden nach validen didaktischen Kriterien strukturiert und analysiert werden. Die Bereitstellung solcher Kriterien war Ziel des dargestellten Kodierleitfadens (Hawelka, 2017).

Dieser Leitfaden kann Evaluatoren dabei unterstützen, bereits während der Erhebung unspezifische studentische Aussagen vor dem Hintergrund didaktischer Kriterien zu klären sowie im Anschluss die Daten nach validen Kriterien zu analysieren und zu interpretieren. Auf dieser Basis und unter Berücksichtigung der Lernziele kann Lehrenden eine strukturierte und zielführende Rückmeldung

zur Steigerung der Lehrqualität gegeben werden. Darüber hinaus ermöglicht eine standardisierte Kodierung auch die Entwicklung verschiedener Referenzrahmen zur Einordnung und zum Vergleich von Lehrveranstaltungen (Hawelka, Hiltmann & Wild, 2017).

Dass einzelne Kategorien nicht immer trennscharf formuliert werden können, ist wesentlich der Komplexität des Gegenstands geschuldet. Dies wurde bereits oben am Beispiel der Kategorie 1 ‚Interaktion‘ skizziert. Gleichzeitig leidet die Trennschärfe darunter, dass Studierende ihre Beobachtungen zum Unterrichtsgeschehen bereits interpretierend rückmelden. Dabei greifen sie keinesfalls auf einheitliche Interpretationsmuster zurück. Studien von Lee, Kim und Chan (2013) belegen, dass Studierende gleiche Aussagen (z. B. ‚Dozent zeigt Enthusiasmus‘) an vollkommen verschiedenen Beobachtungen festmachen (z. B. an rhetorischen Fähigkeiten, dem Einsatz von Motivationsstrategien oder Humor etc.).

Die Kenntnis solcher Hintergründe ist wichtig und muss in die didaktische Interpretation der Daten und in die Lehrberatung mit einfließen. Dies belegt, dass die ‚Teaching Analysis Poll‘ zwar eine im Aufwand überschaubare Evaluationsmethode ist, die aber sowohl bei der Datenerhebung als auch bei der Analyse und Interpretation umfassende didaktische Kenntnisse erfordert, wenn die Ergebnisse in eine fundierte Lehrberatung münden sollen.

7. Literatur

- Biggs, J. & Tang, C. (2011). *Teaching for quality learning at university. What the student does* (4th ed.). Maidenhead: Society for Research into Higher Education and Open University Press.
- Ernst, S. (2008). *Manual Lehrevaluation*. Wiesbaden: Springer VS.
- Frank, A., Fröhlich, M. & Lahm, S. (2011). Zwischenauswertung im Semester: Lehrveranstaltungen gemeinsam verändern. *Zeitschrift für Hochschulentwicklung*, 6 (3), 310–318.
- Hasselhorn, M. & Labuhn, A. (2017). Lernstrategien, metakognitive. In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (18. Aufl., S. 1021). Bern: Hogrefe.
- Hawelka, B. (2017). *Handreichung zur Kodierung qualitativer Evaluationsdaten aus Teaching Analysis Poll* (Schriftenreihe Nr. 5, überarbeitete Version vom März 2017). Regensburg: Universität, Zentrum für Hochschul- und Wissenschaftsdidaktik. Zugriff am 29.11.2017 unter <https://epub.uni-regensburg.de/35379/>
- Hawelka, B., Hiltmann, S. & Wild, K.-P. (2017, März). *Bewertungsmaßstab für Teaching Analysis Poll (TAP). Ein Referenzrahmen zur Rückmeldung qualitativer Evaluationsdaten*. Poster auf der 46. Jahrestagung der Deutschen Gesellschaft für Hochschuldidaktik e. V., Köln.
- Hawelka, B. & Wild, K.-P. (2014, März). *Teaching Analysis Poll (TAP) als Methode der Lehrveranstaltungsevaluation*. Vortrag auf der 43. Jahrestagung der Deutschen Gesellschaft für Hochschuldidaktik e. V., Braunschweig.
- Koch, E. (2004). *Gute Hochschullehre. Theoriebezogene Herleitung und empirische Erfassung relevanter Lehraspekte*. Hamburg: Dr. Kovač.
- Kromrey, H. (2001). Studierendenbefragungen als Evaluation der Lehre? Anforderungen an Methodik und Design. In U. Engel (Hrsg.), *Hochschulranking. Zur Qualitätsbewertung von Studium und Lehre* (S. 11–47). Frankfurt am Main: Campus.
- Lee, H. H., Kim, G. M. L. & Chan, L. L. (2013). Good teaching: what matters to university students. *Asia Pacific Journal of Education*, 35, 98–110.
- Loos, P. & Schäffer, B. (2001). *Das Gruppendiskussionsverfahren. Theoretische Grundlagen und empirische Anwendung*. Wiesbaden: Springer VS.
- Marsh, H. W. (1982). SEEQ: a reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*, 52, 77–95.
- Meyer, H. (2016). *Was ist guter Unterricht?* Berlin: Cornelsen.
- Morgan, D. L. (2009). *Focus groups as qualitative research*. Thousand Oaks, CA: Sage.
- Müller, F. H. (2007). Studierende motivieren. In B. Hawelka, M. Hammerl & H. Gruber (Hrsg.), *Förderung von Kompetenzen in der Hochschullehre* (S. 31–43). Kröning: Asanger.

- Penny, A. R. & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: a meta-analysis. *Review of Educational Research*, 74, 215–253.
- Renkl, A. (2015). Drei Dogmen guten Lehrens: Warum sie falsch sind. *Psychologische Rundschau*, 66, 211–220.
- Rindermann, H. (2003). Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. *Zeitschrift für Evaluation*, 2 (2), 233–256.
- Rindermann, H. (2009). *Lehrevaluation – Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Mit einem Beitrag zur Evaluation computerbasierten Unterrichts*. Landau: Empirische Pädagogik.
- Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen*. Münster: Waxmann.
- Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, 185–200.
- Wolbring, T. (2013). *Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität*. Frankfurt am Main: Campus.

ⁱ Die beiden Skalen ‚Referatsbetreuung‘ (FRADOV) und ‚Referate‘ (HILVE) wurden nicht berücksichtigt, da deren Items keine inhaltlich homogene Komponente abbilden, sondern auf die Bewertung eines speziellen Lehrformats (des Referateseminars) abzielen.