

# Bewertungsmaßstab für Teaching Analysis Poll (TAP). Ein Referenzrahmen zur Rückmeldung qualitativer Evaluationsdaten

Birgit Hawelka, Stephanie Hiltmann & Klaus-Peter Wild

Teaching Analysis Poll (TAP) als Methode zur qualitativen Zwischenevaluation erwies sich zur Fortentwicklung von Lehrveranstaltungen als sehr gut geeignet.

## Ablauf von TAP

Die Lehrperson beendet die Lehrveranstaltung 20 Minuten früher und verlässt den Raum. Ein externer Moderator befragt die Studierenden nach hilfreichen und hinderlichen Aspekten der Lehrveranstaltung. In Kleingruppen diskutieren die Studierenden die Fragen und halten ihre Ergebnisse schriftlich fest. Das studentische Feedback wird der Lehrperson durch den Moderator anonymisiert rückgemeldet (siehe Abbildung 1). Darauf aufbauend entwickeln Berater und Lehrperson gemeinsam Verbesserungsstrategien.

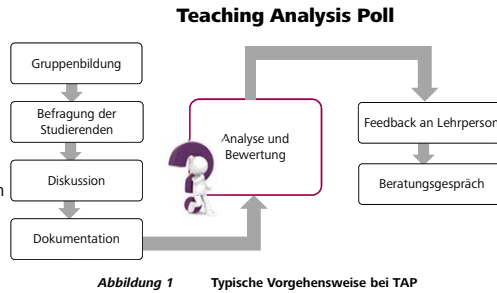


Abbildung 1 Typische Vorgehensweise bei TAP

## Vorzüge von TAP

Konsequente Verknüpfung von Evaluation und Beratung ist entscheidend für die Effektivität der Evaluation. Hohe Effektstärken bei *consultive feedback*  $d_z = 0.69$  (Penny & Coe, 2004)

## Problematik

Bislang keine systematische Analyse und Bewertung der studentischen Rückmeldung unter didaktischen Kriterien  
 ⇒ Keine standardisierte Analyse  
 ⇒ Gewichtung der Rückmeldungen abhängig von der Person des Moderators  
 ⇒ Kein Referenzrahmen zur Bewertung der Rückmeldungen

## Entwicklung eines Kodierleitfadens und eines Kategoriensystems

Grundlage zur Entwicklung des Kategoriensystems war die Triangulation einer theoretischen Beschreibung von Lernprozessen und empirisch überprüfter Aspekte guten Lehrverhaltens. Zur Analyse der Evidenzen guter Lehre wurden drei standardisierte Evaluationsinstrumentarien herangezogen: SEEQ (Marsh, 1982), HiLVE (Rindermann, 2009) und FRADOV (Koch, 2004). Die Instrumentarien wurden zunächst auf Skalenebene inhaltlich analysiert und separiert (siehe Abbildung 2).

Anschließend wurden die Skalen zum Bereich *Lernen* gemäß ihrer didaktischen Funktion zur Unterstützung studentischer Lernprozesse (Wild, 2000) weiter differenziert sortiert (siehe Abbildung 3). Aus diesem Modell wurde ein Schema zur Kategorisierung qualitativer Daten abgeleitet und ein Kodierleitfaden mit Definition der Kategorien und Ankerbeispielen erstellt (Hawelka, 2016, Kodierleitfaden online erhältlich unter [www.uni-r.de/zhw/forschung](http://www.uni-r.de/zhw/forschung)).

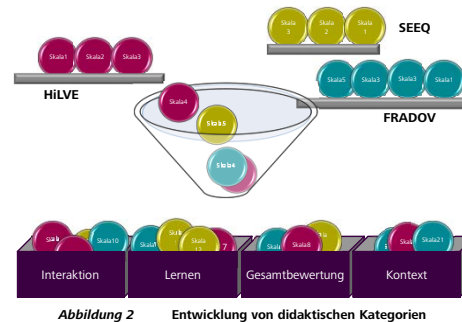


Abbildung 2 Entwicklung von didaktischen Kategorien

## Validität

- Konstruktvalidität durch Anlehnung an valide Evaluationsbögen
- Kommunikative Validierung während der Diskussion und Dokumentation

## Reliabilität

- 58 TAPs (1772 Kodiereinheiten) aus 10 Fakultäten
- Intercoderreliabilität (UK) Cohen's  $k = .833$  (HK) Cohen's  $k = .865$

## Vollständigkeit

- Zuordenbare Items: 98,9

(Hawelka & Wild, 2016)

## Stichprobe

53 TAP von Seminaren der philosophischen Fakultät an der Universität Regensburg  
 978 Studierende  
 195 Gruppen  
 1525 Aussagen

## Interpretation

Studierende der philosophischen Fakultäten geben mehr positives als negatives Feedback. Offensichtlich wichtig aus Sicht der Studierenden sind die Kategorien

- *Verständliche & anregende Darstellung*
- *Aktive Beteiligung & Interaktionsmanagement*
- *Lehre planen & strukturieren*
- *Lernmaterialien / Literatur*

da sie in den meisten TAP positiv wie negativ genannt wurden.

Einige Kategorien werden vorwiegend genannt, wenn sie positiv auffallen, jedoch nicht negativ bewertet, wenn sie im Lehrverhalten nicht sichtbar sind:

- *Soziale Eingebundenheit*
- *Interesse & Engagement des Lehrenden*
- *Inhalte kritisch prüfen*
- *Adaptiv regulieren*
- *Ansprechpartner bei Problemen.*

Im Gegensatz dazu weisen die Ergebnisse

darauf hin, dass die Kategorien

- *Veranstaltung leiten*
- *Erwartungsklarheit*
- *Kompetenzerleben*

hauptsächlich dann wahrgenommen werden, wenn Lehrende sie nicht ausreichend erfüllen.

## Ergebnisse

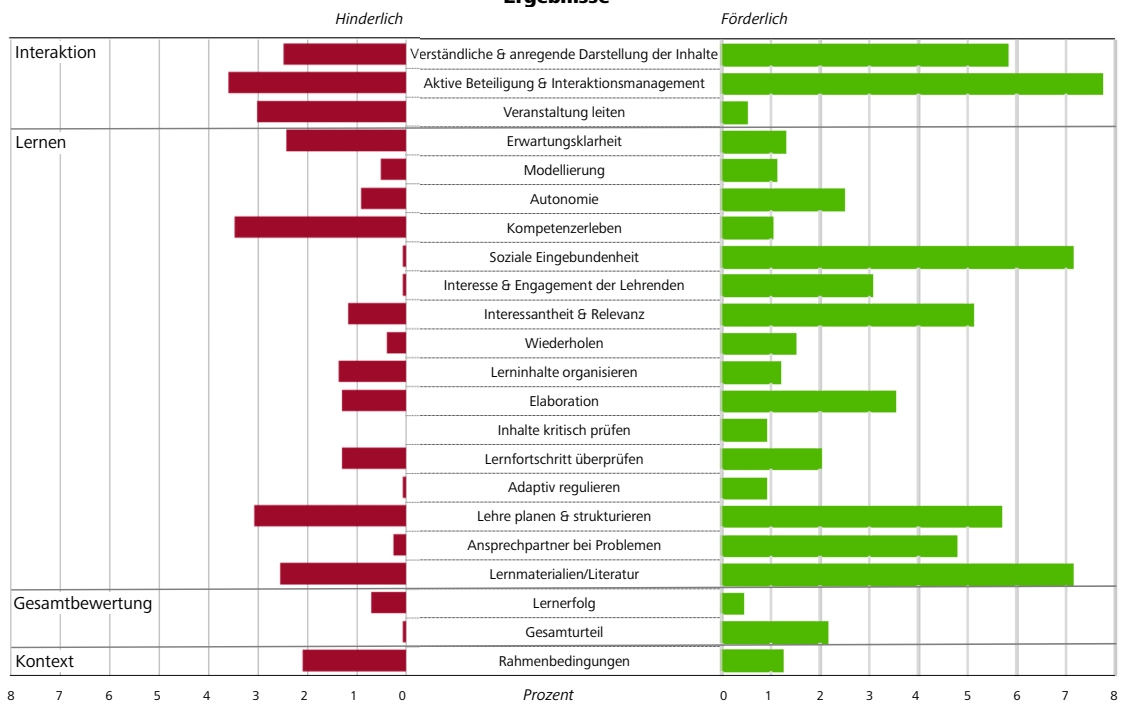


Abbildung 3 Referenzrahmen für die philosophischen Fakultäten: Prozentuale Häufigkeiten der Aussagen verteilt auf die Kategorien

## Ausblick

Zukünftige TAP-Ergebnisse von Lehrveranstaltungen aus philosophischen Fakultäten können nun mit dem Referenzrahmen abgeglichen werden, um dadurch Besonderheiten und spezifische Probleme zu identifizieren. In weiteren Forschungsarbeiten soll ein Referenzrahmen für naturwissenschaftliche Fakultäten entwickelt und mit den hier präsentierten Ergebnissen hinsichtlich der prozentualen Häufigkeitsverteilungen und Gewichtungen der Kategorien verglichen werden.

## Literatur

- Frank, A., Fröhlich, M. & Lahn, S. (2011). Zwischenauswertung im Semester: Lehrveranstaltungen gemeinsam verändern. *Zeitschrift für Hochschulentwicklung*, 6 (3), 310-318.
- Hawelka, B. (2016). *Handreichung zur Kodierung qualitativer Evaluationsdaten aus Teaching Analysis Poll*. Schriftenreihe (ZHW) 5, Universität Regensburg, Regensburg.
- Hawelka, B. & Wild, K.-P. (2016, September). *Auswertung von Teaching Analysis Poll (TAP). Ein Kodierleitfaden zur Analyse qualitativer Evaluationsdaten*. Beitrag zur Jahrestagung der Deutschen Gesellschaft für Hochschuldidaktik, Bochum.
- Koch, E. (2004). *Gute Hochschullehre: Theoriebezogene Herleitung und empirische Erfassung relevanter Lehraspekte*. Hamburg: Kovač.
- 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.
- Penny, A. & Coe, R. (2004). Effectiveness of Consultation on Student Ratings Feedback: A Meta-Analysis. *Review of Educational Research*, 74 (2), 215-253.
- Rindermann, H. (2009). *Lehrevaluation - Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Mit einem Beitrag zur Evaluation computerbasierter Unterrichts*. Landau: Empirische Pädagogik.
- Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen*. Münster: Waxmann.