

Univ.-Prof. Dr. Heiner Rindermann
Universität Graz

**Evaluation von Lehrveranstaltungen -
eine Chance zur Weiterentwicklung für
alle Beteiligten
(Studierende, Lehrende, Rektorat)**

1 Definition und Beschreibung guter Lehre als Qualitätsmerkmal

Ansätze zur Bestimmung von Kriterien guter Lehre

- Definition von Lehrzielen für Lehre allgemein oder für Lehrveranstaltungen
- Orientierung an Gesetzestexten, Bildungsidealen, Curricula, Lehrevaluationsbögen
- Heranziehung von hochschuldidaktischen und instruktionspsychologischen Theorien und Forschungsergebnissen

- Beobachtung und Beschreibung der Veranstaltungen von als „guten“ Lehrkräften bekannten Personen
- Prädiktoren von Lehrerfolg oder allgemeiner Dozentenbeurteilung
- Offene Befragung von Lehrenden und Studierenden
- Relevanzrating vorgegebener Items

Empirische Studie (Rindermann, 1999):

Befragung von 58 Studierenden und 58 Lehrenden verschiedener Fächer und Hochschulen nach Merkmalen guter Lehrkräfte und Lehre:

Merkmale guter Lehrkräfte:

Didaktische Fertigkeiten, z. B. Strukturierung, adäquate Themenauswahl und Anforderungsanpassung,

zusätzlich bedeutsam *Persönlichkeitsmerkmale* wie Freundlichkeit,

soziale Verhaltensweisen wie Kooperativität und Motivierung

Merkmale guter Veranstaltungen:

Gute Veranstaltungen werden vor allem über **gutes Handeln der Lehrenden** definiert.

Häufiger werden hier **Themenauswahl**, **Anforderungsniveau** und **Interaktion** angeführt.

Übereinstimmung zwischen Studierenden und Lehrenden $r = .79$
(Berechnung über Items)

Lehrende betonen aber eher fachlich-inhaltliche Aspekte,
Studierende die Vermittlung von Inhalten.

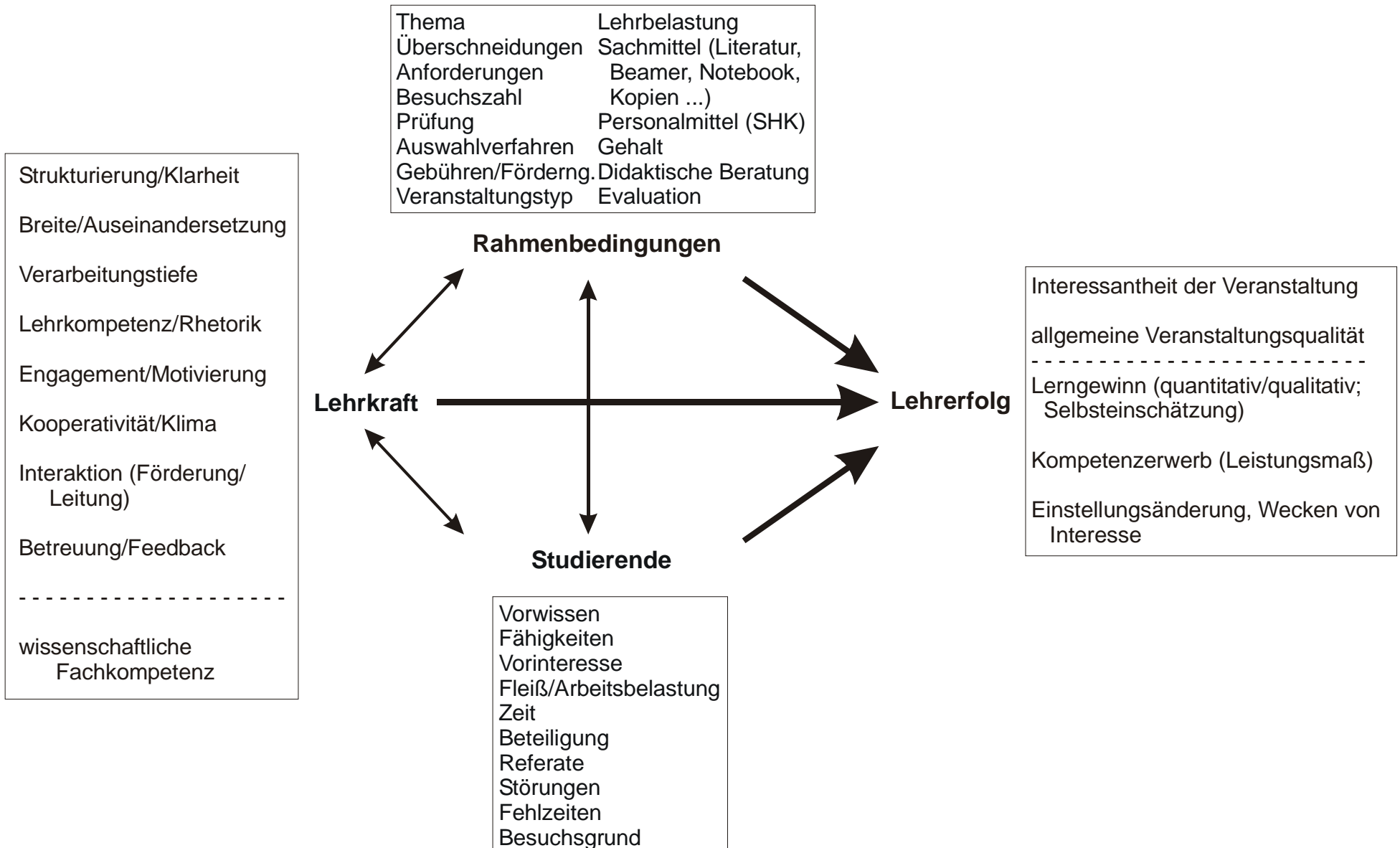
***** Exkurs Korrelation *****

Korrelationen geben die Enge und Richtung eines Zusammenhangs an.
Korrelationskoeffizienten (r) variieren zwischen -1 , 0 und $+1$.

Z. B. $r=.70$ zwischen Körpergröße und -gewicht oder $r=-.40$ zwischen Hai-dichte und Schwimmerzahl.

Bei $r=\pm.10$ spricht man von einem schwachen, bei $r=\pm.30$ von einem mittleren und ab $r=\pm.50$ von einem starken Zusammenhang (Cohen, 1988).

Korrelationen erhalten erst durch eine angemessene inhaltliche Deutung an Relevanz.



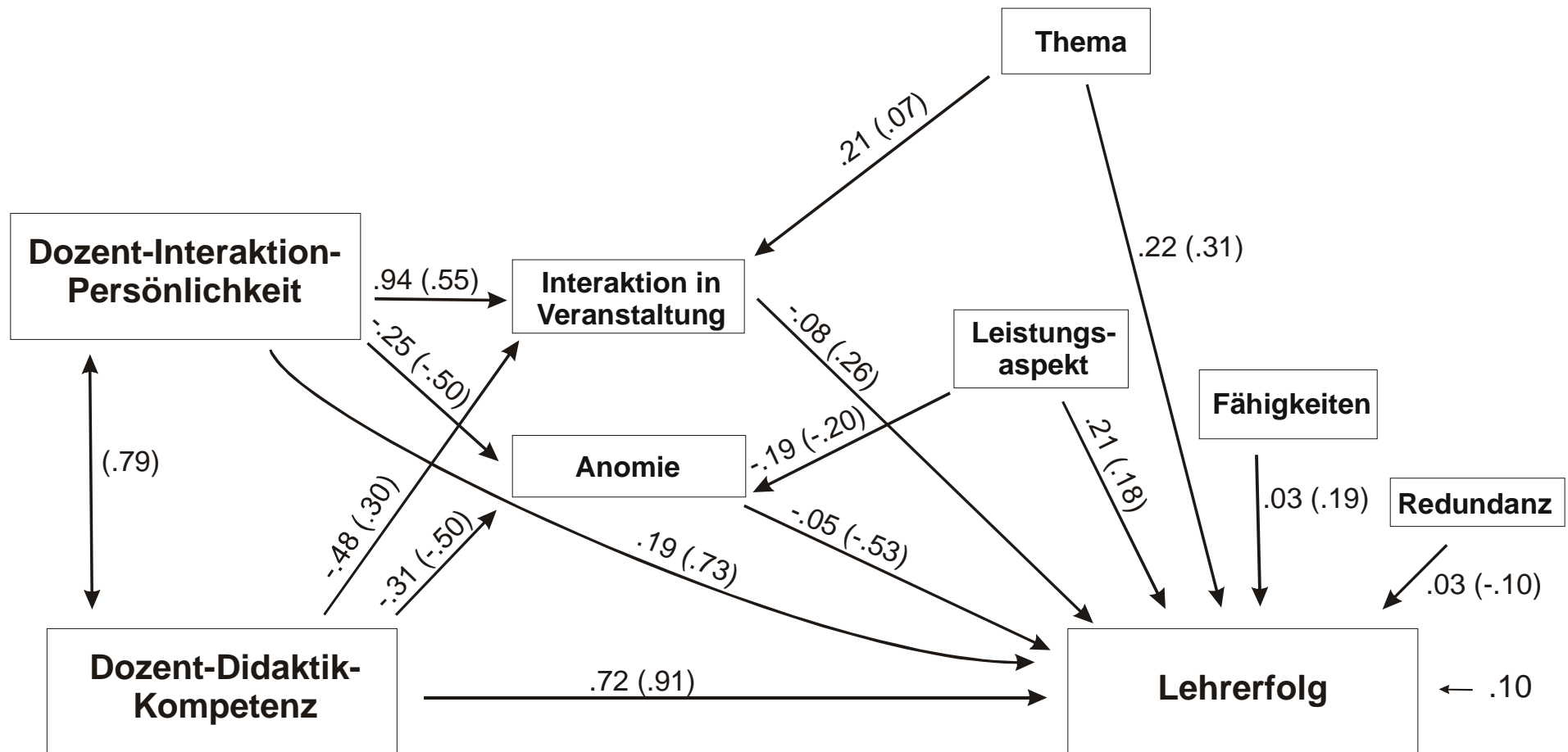


Abbildung: Strukturgleichungsmodell für HILVE-II, hier für Vorlesungen (standardisierte Koeffizienten, in Klammern Korrelationen), N=73 Veranstaltungen

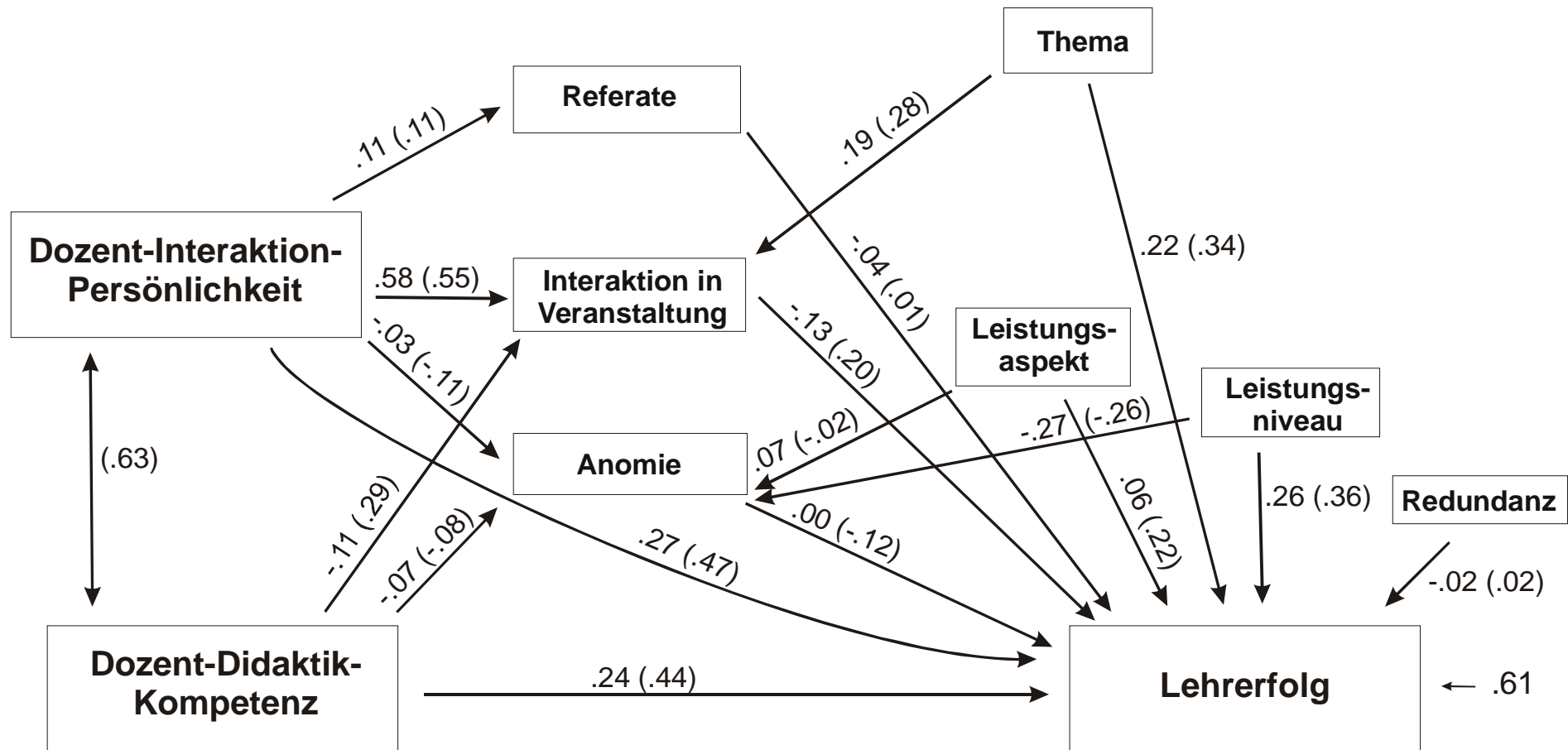


Abbildung: Strukturgleichungsmodell für HILVE-II, hier für Einschätzungen durch Lehrende, N=154

2 Konstruktion von Instrumenten zur Beurteilung der Lehre

Ziel:

- Fragebogen, der inhaltlich relevante Aspekte der Lehre erhebt
- Fragebogen, der den Qualitätsanforderungen an Testverfahren Genüge leistet
- Fragebogen, der im praktischen Einsatz akzeptiert wird
- Fragebogen, dessen Einsatz und Verwendung intendierte Effekte zeigt

Optimum: Instrumente,

- die aus einem Basisitemsatz und fakultativen, auf die jeweilige Veranstaltung adaptierten Itembausteinen bestehen
- die sowohl geschlossene, standardisierte wie auch offene, freie Items beinhalten
- die multidimensional aufgebaut sind
- deren Items und Skalen hinsichtlich der Testgütekriterien überprüft worden sind
- und mit Berücksichtigung der Erkenntnisinteressen der Nutzer, den Studierenden und Lehrenden, entwickelt wurden.

Itemgenerierung:

- *Befragung*
- *Synkretistisch*
- *Gremienmodell*
- *Theoretisch*
- *Autor/in*

Itemselektion nach

- inhaltlicher Breite, Relevanz und Angemessenheit
- Verständlichkeit und Kürze
- Objektivität, Reliabilität, Validität
- Faktorenstruktur
- Feedbackrelevanz, Nützlichkeit
- Ökonomie
- fairen Fragen

3 Verbesserung von Lehrqualität durch Evaluation?

Systemisch-gesellschaftliche Ebene: Bargel, Multrus und Ramm (2005)

Säkularer Trend in Deutschland: Verbesserung der Lehre in Befragungen von Studierenden zu Lehrveranstaltungen und Lehre allgemein

Lehrveranstaltungen

(Prozentangaben: Anzahl aller befragten Studierenden bundesweit mit gut)

- Vortrag ist verständlich und treffend, 1993 → 2004, 35% → 53%: **+18%**
- Lehrende sind gut vorbereitet, 1993 → 2004, 28% → 42%: **+14%**

Lehre allgemein

(Prozentangaben: Anzahl aller befragten Studierenden bundesweit mit gut)

- Lehrveranstaltungen, 1983 → 2004, 28% → 54%: **+26%**
- Beratung und Betreuung, 1987 → 2004, 24% → 41%: **+17%**

**Lehrevaluation soll nicht nur diagnostizieren, sondern helfen,
Lehre zu verbessern.**

Überprüfung der Effektivität von Lehrevaluation als solcher oder Lehr-
evaluation und mit ihr verbundener Maßnahmen durch wiederholte Messun-
gen innerhalb und zwischen Semestern:

Lehre - Lehrevaluation (+Treatment) - Lehre - Lehrevaluation

optimal: mit Vergleichsgruppen

Literaturstand:

- Sensibilisierungshypothese 0-Effekt
- Feedback-Hypothese $d=0,2$
- hochschuldidaktisches Diskursmodell 0-Effekt
- Beratungs-Trainings-Ansatz $d=0,64$ (P. A. Cohen, 1980)
- Beratungs/Trainings/Gratifikationsansatz ?

***** Exkurs d *****

Mittelwertunterschiede werden vergleichbar mit d dargestellt.

Nach Cohen (1988) gelten Differenzen im Effektstärkemaß d von $d=\pm 0,2$ als klein, von $\pm 0,5$ als mittel und von $\pm 0,8$ als groß.

Eigene Studien: Feedbackstudie an einer Universität

Hintergrund

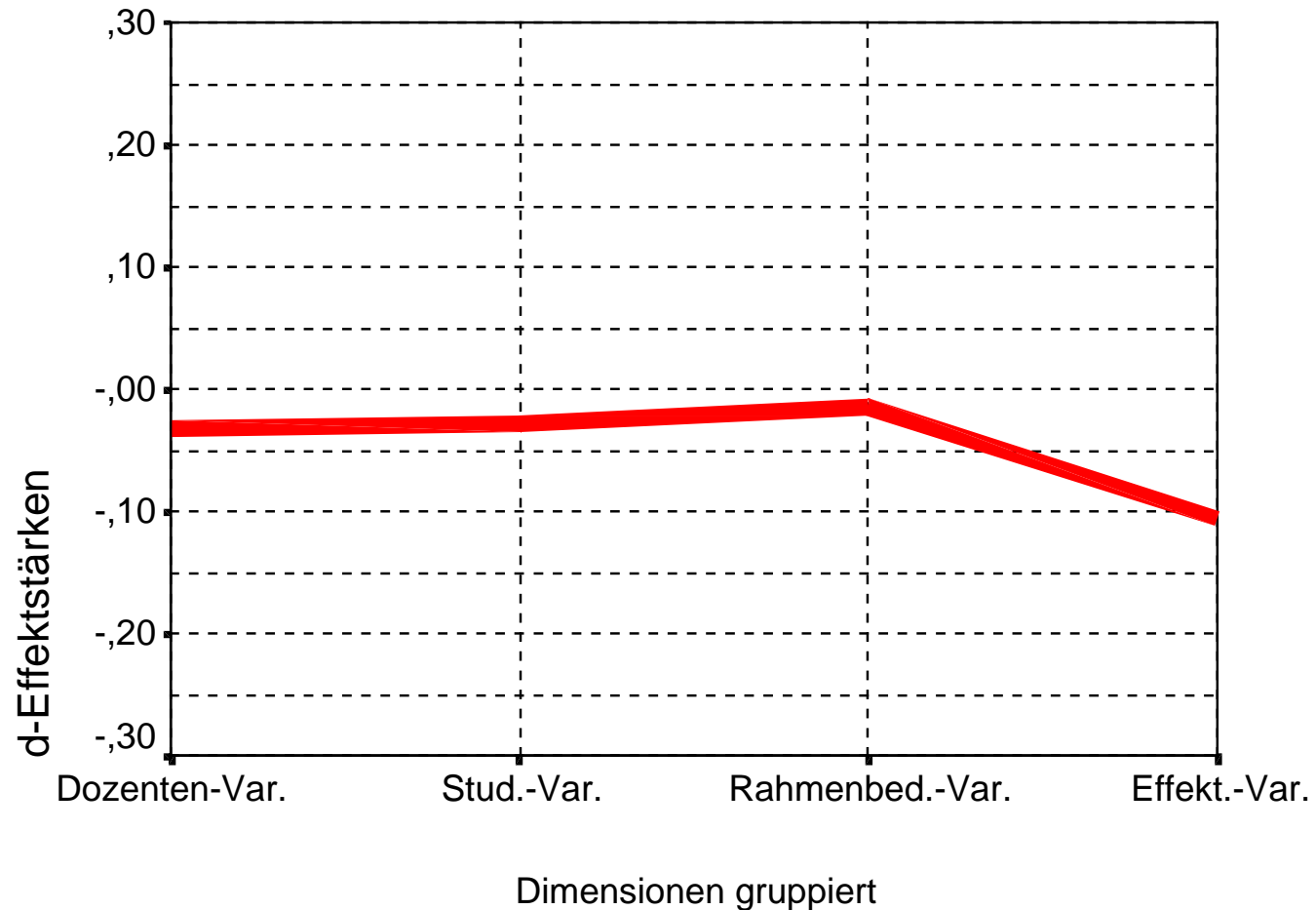
- Erhebungszeitraum Anfang bis Mitte der 90er Jahre
- Sozial- und Geisteswissenschaften
- wiederholt bewertet durch Studierende, Lehrende und Fremdrater (HILVE-I)
- für alle Veranstaltungen Feedback
- 42 Veranstaltungen mit und ohne Besprechung der Ergebnisse mit Studierenden
- Änderungen innerhalb eines Semesters (Mitte vs. Ende, nur identische oder auch verschiedene Studierende) und über verschiedene Semester hinweg untersucht
- Rindermann (1996, S. 146ff.)

Ergebnisse

$d \approx 0,0$

Messzeitpunktveränderung in HILVE-I-Skalen

universitäre Feedbackstudie



Konklusion

Feedback sowie Besprechung und Lehrevaluation als solche sind nicht effektiv.

Beratungsstudie an privater Ausbildungsschule für Logopädie

Hintergrund

- Erhebungszeitraum Ende der 90er Jahre bis 2002
- Logopädie (Schüler/innen: Erwachsene, viele mit Vorqualifikationen)
- 35 Kursbewertungen durch Veranstaltungsteilnehmende (HILVE-II)
- Rückmeldung + Beratung, Institutskonferenzen
- Änderungen über verschiedene Semester und Schülerjahrgänge hinweg
- Rindermann & Kohler (2003), Rindermann, Kohler & Meisenberg (2007)

Bestandteile der Rückmeldung und Beratung

- Erläuterung des statistischen Feedbacks.
- Diskussion der Selbstsicht der Lehrkraft im Vergleich mit den studentischen Angaben.
- Eruierung von Gründen für positive oder negative Ergebnisse.
- Besprechung einzelner Unterrichtsformen und didaktischer Strategien.
- Erarbeitung von ca. *drei bis fünf konkreten Vorschlägen* zur Veränderung des Handelns im Unterricht.
- Zusammenfassung, Signalisierung von jederzeitiger Gesprächsbereitschaft und Interesse des Beraters an der weiteren Entwicklung der Lehrkraft.

Beratung dient als

- **Information** (Feedbackerläuterung),
- **Anregung von Reflexion und Motivation**,
- **Intervention** (Stützung und Aufbau neuer Lehrtechniken und pädagogischer Handlungsweisen) und
- **Normfokussierung** (gute Lehre ist wichtig).

Auf Ebene der Institution:

Konferenzen von Lehrenden und Lernenden.

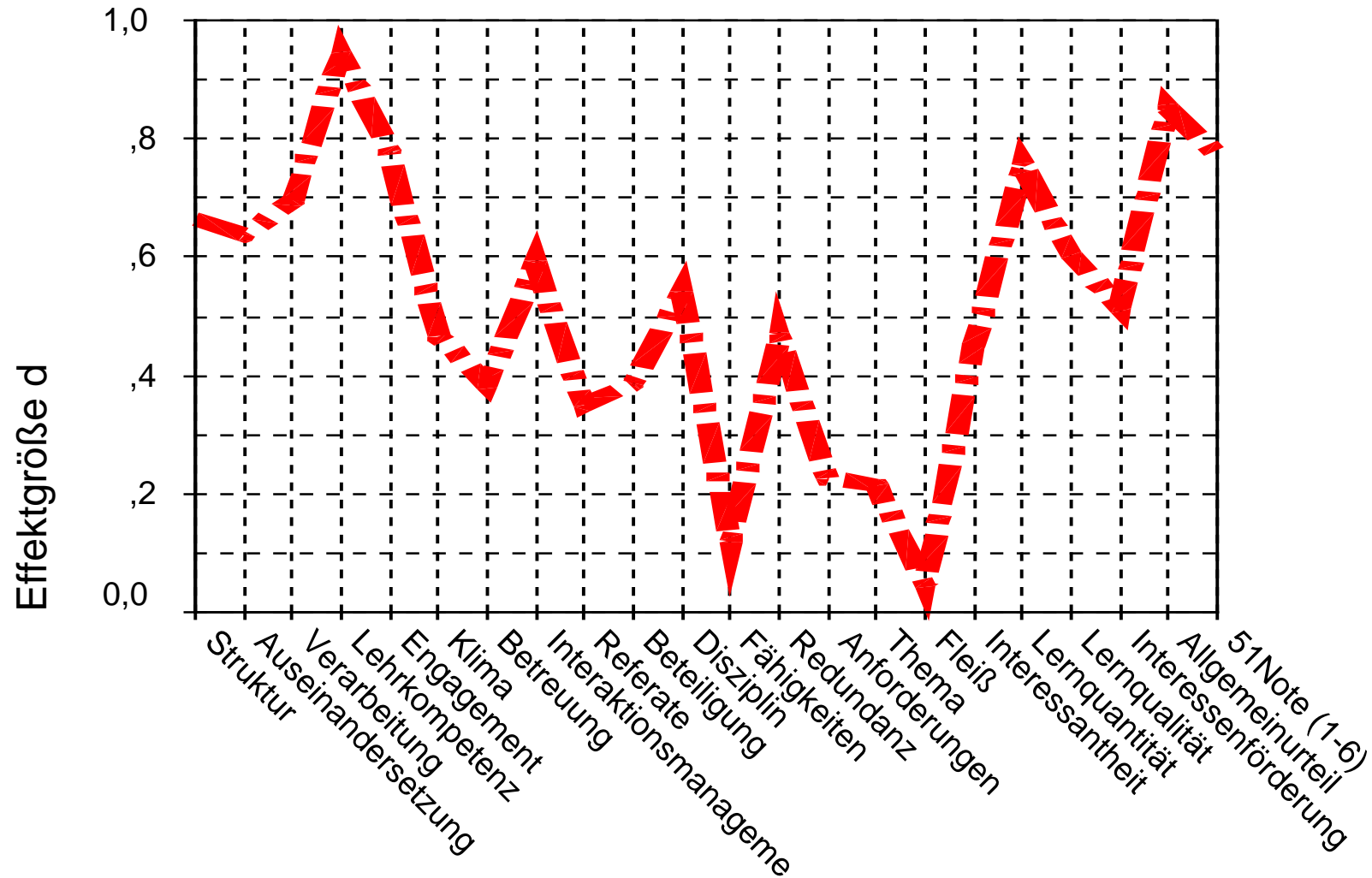
Ergebnisse

	Dozenten- skalen	Lehr- effektivität	studentische Skalen	Rahmen- bedingungen
d_{MW}	0,65	0,64	0,43	0,24

Lehrkompetenz: **0,94**; Dozentenengagement: **0,78**; Allgemeinurteil: 0,85

Messzeitpunktveränderung in HILVE-II-Skalen

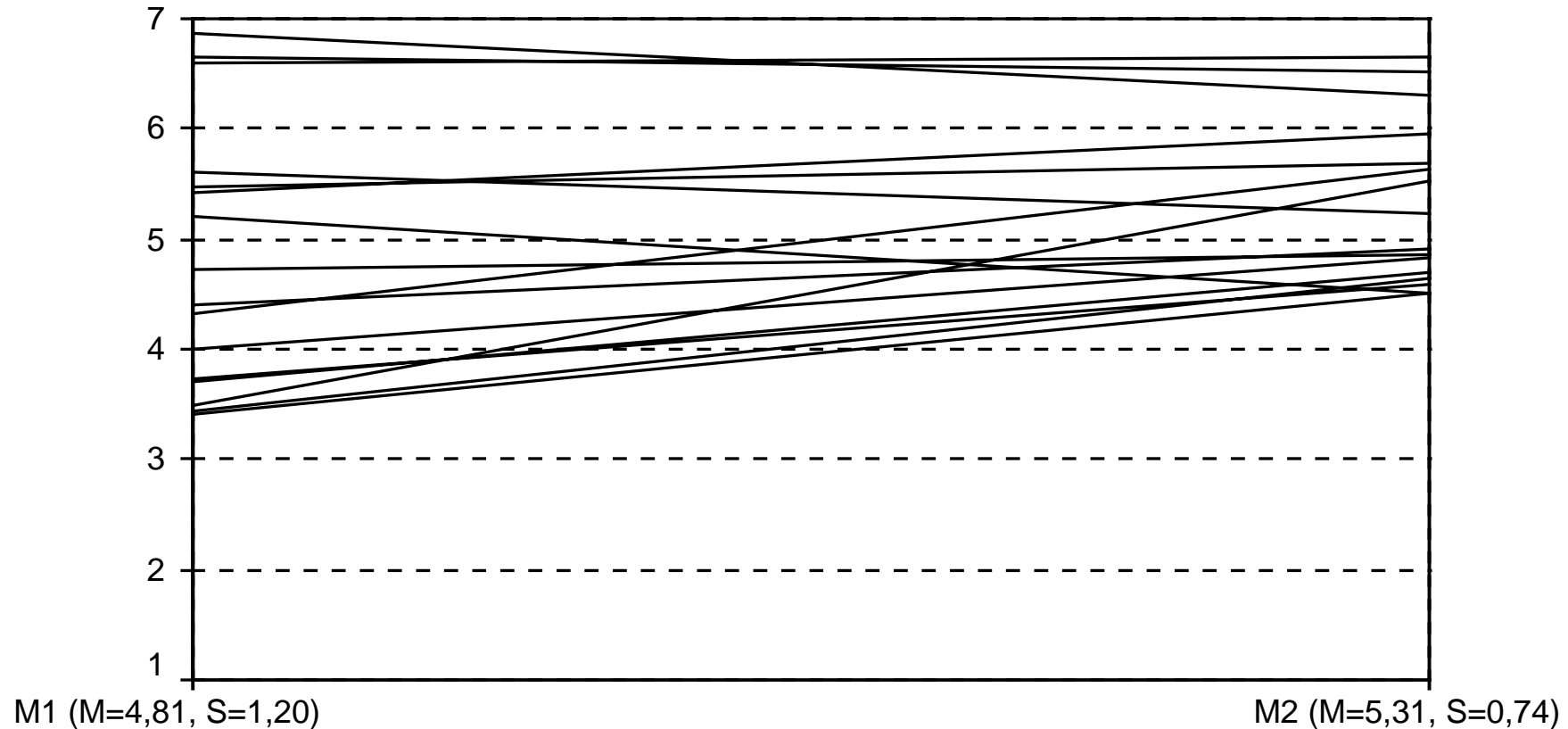
16 Lehrkräfte mit 35 Veranstaltungen



d-Effektgrößen für alle Skalen und die Veranstaltungsnote

Messzeitpunktveränderung in Lehrkompetenz

16 Lehrende mit 35 Veranstaltungen



Rohwerte (Skala 1 bis 7, hohe Werte positiv)

Messung 1 - Messung 2, je Linie eine Lehrkraft

Mittelwertsverbesserung und Streuungsreduktion in der Lehrkompetenz ($r_{tt}=.87$ in Dozentenskalen)

Konklusion

Lehrevaluation mit Beratung (und bei institutioneller Wertschätzung von Lehrqualität) führt zu deutlich erkennbaren Verbesserungen.

Beratungsstudie an einer staatlichen Fachhochschule

Hintergrund

- Erhebungszeitraum Beginn-Mitte des ersten Jahrzehnts nach 2000
- Studiengang Architektur an einer Fachhochschule (18 SWS)
- studentische Einschätzungen (HILVE-II) für 12 Lehrkräfte in 98 Kursen
- Rückmeldung + Beratung
- Änderungen über verschiedene Semester (zwei Jahre Abstand) und Studierende hinweg: mittelfristige Veränderungen und Transfer
- Dresel, Rindermann & Tinsner (2007)

Methodik

für die Berechnung der *Veränderungswerte*

Beachtung und Herausrechnung von möglichen *Verzerrungsfaktoren* (individuelles Vorwissen, Themenrelevanz, Geschlecht, Vorinteresse, Anforderungshöhe; Vorinteressenmittel, Vorwissensheterogenität; Dozentenmerkmale)

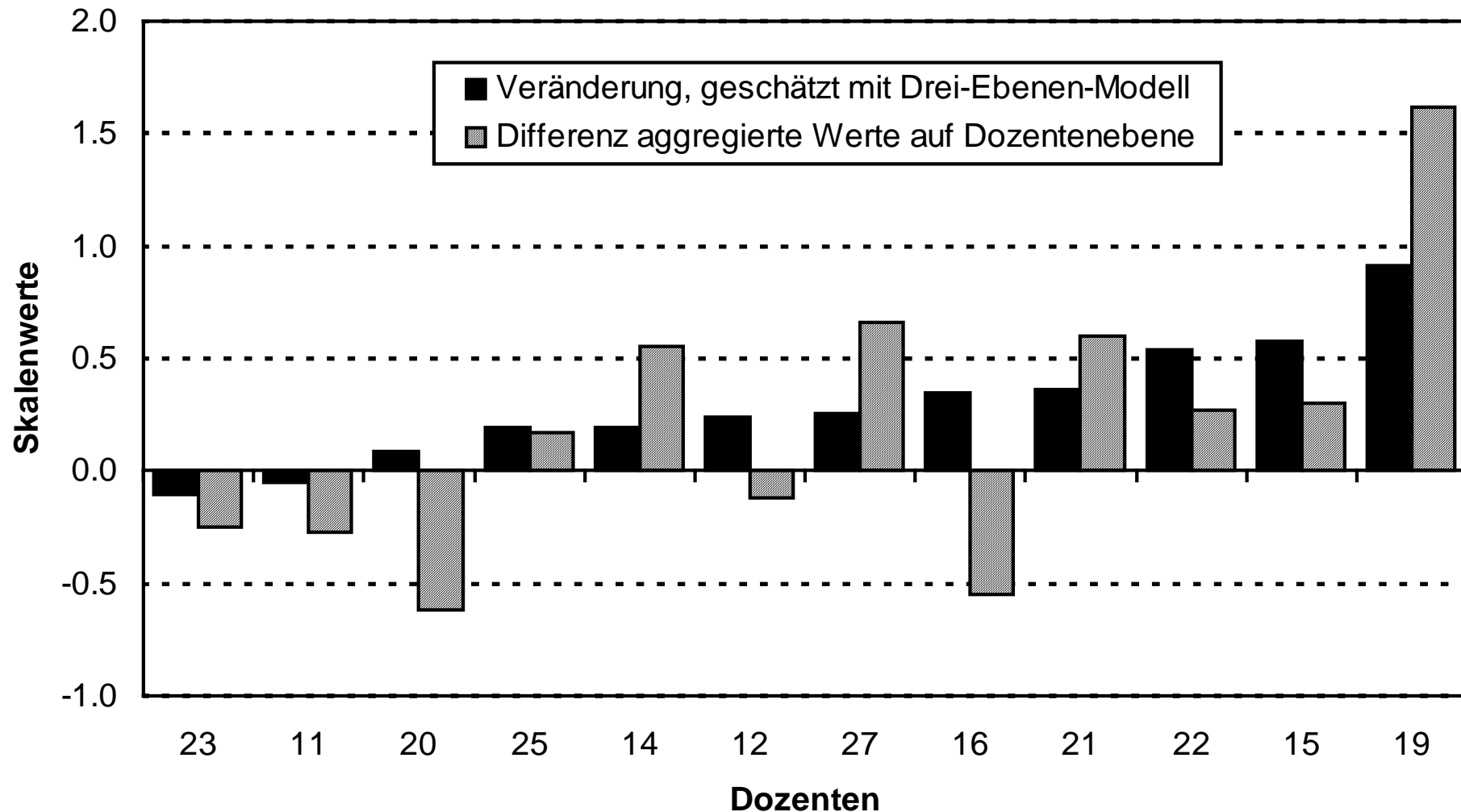
auf *drei Ebenen* (Studierende, Veranstaltungen, Lehrende)

→ So ist Vergleichbarkeit über verschiedene studentische Gruppen und Veranstaltungen verbessert.

Ergebnisse

	Lehreffektivität (Lernen quantitativ und qualitativ, Interessenförderung und Allgemeinbeurteilung)
d_{MW}	0,68 - 0,91

(Verbesserung des Lehrerfolgs der beratenen Lehrkräfte, berechnet mit einem Drei-Ebenen-Veränderungsmodell unter Berücksichtigung von Biasvariablen)



Veränderung der beurteilten Lehreffektivität: korrigierte (schwarze) und unkorrigierte (grau-schraffierte) Werte für 12 Lehrkräfte (Korrelation zwischen beiden $r = .79$)

Konklusion:

Lehrevaluation mit Feedback und Beratung ist auch an staatlichen Hochschulen effektiv.

Aber:

Problem: Es gibt kaum Beratung.

4 Besonderheiten der HILVE-Anwendung (unter EvaSys)

HILVE,

Heidelberger Inventar zur Lehrveranstaltungsevaluation:

- gemäß gängiger testdiagnostischer Kriterien entwickeltes und überprüfbares Instrument
- seit 1994 im Einsatz, seit 1999 revidiert
- hohe thematische Breite, multidimensional, Lehreffektivitätsmodell
- Normen
- Besonderheit: adaptives Beratungsangebot per Mail auf Wunsch (Kompromiss zwischen wünschenswerterer persönlicher Rückmeldung und Beratung und bisheriger Praxis der Nichtberatung)

5 Empfehlungen zur Durchführung von Lehrevaluation als Qualitätssicherung

- Verwendung eines Skalen- und offene-Fragen-Verfahrens
- vorgegebene Items, Skalen und Antwortstufen
- offene Fragen: gut/schlecht/Anregungen zu Verbesserungen

- Evaluation alle drei Semester aller Veranstaltungen eines Fachbereichs

- nicht nur Lehrveranstaltungen, sondern auch Lehre im Allgemeinen evaluieren (sachliche Ausstattung, Bibliothek, numerische Relation Lehrende-Studierende, Sprechstunden, Betreuung, Studiendauer)

- Feedback (Normen!), Beratung und hochschuldidaktische Trainingsangebote: *Beratungs-Rückmelde-Verfahren*
- Oder Fragebogen mit automatisierten Rückmeldetexten je nach Ergebnis (Zwischenstufe vor persönlicher Beratung und nach bloßem Ergebnisfeedback)
- Nicht: Evaluation - die bloße Messung der Lehrqualität - von oben zu verordnen, ohne Personen und Institutionen zu ermöglichen, Qualität zu verbessern.
- Einbau der Lehrevaluation in ein Qualitätssicherungsmodell von Lehre und Forschung
- Individuelle Qualität gibt auch strukturelle Organisationsqualität wieder.
- Hochschulautonomie (Entscheidungsmöglichkeiten)

Literaturvorschläge

- Dresel, M., Rindermann, H. & Tinsner, K. (2007). Beratung von Lehrenden. In A. Kluge & K. Schüler (Hrsg.), *Qualitätssicherung und -entwicklung an Hochschulen* (S. 193-204). Lengerich: Pabst.
- Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Verlag für Psychologie.
- Mummendey, H. D. & Grau, I. (2007). *Die Fragebogen-Methode*. Göttingen: Hogrefe.
- Rindermann, H. (2003). Lehrevaluation an Hochschulen. *Zeitschrift für Evaluation*, 3(2), 233-256.
- Rindermann, H. (2009). *Lehrevaluation - Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen*. Landau: Empirische Pädagogik.
- Rindermann, H. (1998). Skalen der Lehrevaluation: Welche Aspekte sollen in universitären Lehrveranstaltungen beurteilt werden? In G. Krampen & H. Zayer (Hg.), *Psychologiedidaktik und Evaluation*. (S. 295-316). Bonn: DPV.