

Wirksamkeit von mehrstufigen Evaluationsverfahren in Studium und Lehre

Netzwerk Quality Audit: Verbundtreffen an der Universität des Saarlandes
Saarbrücken, 24. September 2020

Autor*innen: Andreas Fritsch, Jessica Walter und Lara Lichtenthäler (Universität Greifswald)

UNIVERSITÄT GREIFSWALD
Wissen lockt. Seit 1456



Einleitung

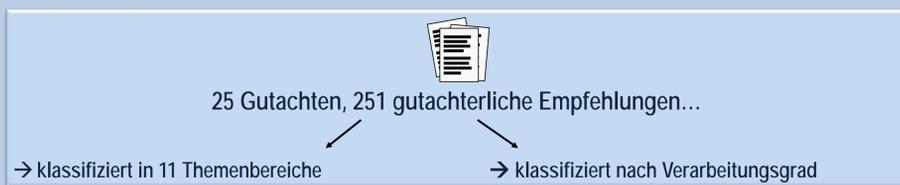
Mehrstufige Evaluationsverfahren (Bornmann/Mittag/Mutz/Daniel, 2004) sind in der Qualitätssicherung von Studium und Lehre weit verbreitet. Mit Abschluss eines Evaluationszyklus von mehrstufigen Evaluationsverfahren, die im Zeitraum 2012-2017 an einer Hochschule durchgeführt wurden, sollte Bilanz gezogen werden zu Ergebnissen und Wirkungen. Mit ihrer Studie Qualitätssicherung an Hochschulen hatten Sandra Mittag (2006) und Kolleg*innen (Mittag/Bornmann/Daniel, 2006) die Wirksamkeit der mehrstufigen Evaluationsverfahren u. a. des Verbundes Norddeutscher Universitäten aus dem Zeitraum 1995-2003 untersucht.

Mit der vorliegenden Untersuchung wird ein mehrstufiges Evaluationsverfahren betrachtet, welches auf den von der Autorin untersuchten Verfahren aufbaut. Daher bot es sich an, dessen Wirksamkeit unter Verwendung der gleichen Untersuchungsmethoden, v. a. Dokumentenanalysen, zu evaluieren. Das hochschulintern organisierte mehrstufige Evaluationsverfahren unterscheidet sich v. a. dadurch von dem Vorgängerverfahren, dass die Begutachtungen und auswertenden Konferenzen zur Umsetzung der gutachterlichen Empfehlungen nicht hochschulübergreifend im Verbund durchgeführt wurden. Die damit verbundenen geringeren Vergleichs- und Lernmöglichkeiten durch Benchmarking oder Modelllernen begründen die Untersuchungshypothese, dass die intern organisierten Evaluationsverfahren eine geringere Wirksamkeit entfalten als Verbundevaluationsverfahren.

Hypothese: Wirksamkeit Verbundevaluationsverfahren > Wirksamkeit interner mehrstufiger Evaluationsverfahren

Methodik

Es wurden die Gutachten und Umsetzungsberichte von 25 mehrstufigen Evaluationsverfahren ausgewertet, die im Zeitraum 2012-2017 an der Hochschule durchgeführt worden waren. Die Dokumentenanalyse explizierte insgesamt 251 gutachterliche Empfehlungen. Gemäß der Systematik von Mittag (2006) wurden die gutachterlichen Empfehlungen in elf Themenbereiche klassifiziert. Aus den Protokollen der auswertenden Veranstaltungen zu den Gutachten wurden die Berichte zur Umsetzung der gutachterlichen Empfehlungen entnommen. Deren Bearbeitungsgrad wurde in drei Stadien unterteilt. Die Kodierung der Dokumente erfolgte durch zwei unabhängige Codierer*innen. Die Beurteilungsübereinstimmung betrug für die Codierung der Themenbereiche der gutachterlichen Empfehlungen ausgezeichnete 97 % (Cohens $k=0.94$) und für die Codierung der Umsetzungsaktivitäten akzeptable 81 % (Cohens $k=0.62$).



Interpretation

Die Ergebnisse scheinen zunächst die Hypothesen zu stützen, dass die intern organisierten Evaluationsverfahren eine geringere Wirksamkeit entfalten als Verbundevaluationsverfahren.

Es fällt auf, dass bei den intern organisierten Evaluationsverfahren vergleichsweise häufig gutachterliche Empfehlungen begründet nicht umgesetzt wurden. Diese Besonderheit wird im Zusammenhang mit den unterschiedlichen Schwerpunktsetzungen der gutachterlichen Empfehlungen in beiden Evaluationsverfahren gesehen: In der Studie von Mittag (2006) betrafen die meisten gutachterlichen Empfehlungen, anders als beim internen Verfahren ($p < 0.001$), die Planung und Organisation von Studium und Lehre (36 %), gefolgt von Empfehlungen zur Ausstattung (17 %) sowie zu Lehr- und Lernformen (8 %). Die wenigsten Empfehlungen waren für Qualitätssicherung und -verbesserung von Studium und Lehre (4 %), Bildungs- und Ausbildungsziele (2 %) sowie Verwaltung und akademische Selbstverwaltung (2 %) formuliert worden (Mittag 2006: 80). Bei den zwischen 1995-2003 durchgeführten mehrstufigen Evaluationsverfahren wurden also grundsätzlich andere thematische Schwerpunkte gesetzt als bei den zwischen 2012 und 2017 organisierten internen Verfahren. Diese basieren auf den Themenbereichen und Bewertungsmaßstäben für die Begutachtung von Studienprogrammen und von qualitätssichernden Verfahren gemäß der „Kriterien für die Akkreditierung von Studiengängen“ (Beschluss des Akkreditierungsrats vom 08.10.2007) und der „Kriterien für die Systemakkreditierung“ (Drs. AR 11/2008).

Die Feststellung der Akkreditierungsfähigkeit spielte bis 2003 keine Rolle bei den gutachterlichen Bewertungen. Die Möglichkeit des Vergleichs der Leistungsfähigkeit der beteiligten Fachbereiche anhand der Gutachten und Umsetzungsberichte stellte sicherlich einen Anreiz für Umsetzungsaktivitäten dar. Bei den internen mehrstufigen Evaluationsverfahren wurden grundsätzlich alle Auflagen im Zusammenhang mit der Akkreditierung von Studienprogrammen umgesetzt. Maßstab für den Erfolg eines Evaluationsverfahrens ist vor allem auch der Erhalt des Siegels des Akkreditierungsrats, weniger der Vergleich mit anderen Fachbereichen.

Des Weiteren fällt auf, dass die Verbindlichkeit der Dokumentation der Umsetzung der gutachterlichen Empfehlungen bei der Verbundevaluation deutlich höher erscheint. Die Fachbereichsleitung verfasste einen Umsetzungsbericht, der zusammen mit den Gutachten und den Stellungnahmen der beteiligten Fachbereiche veröffentlicht wurde. Zwischen Gutachten und Umsetzungsbericht lagen rund eineinhalb Jahre. Der Umsetzungsbericht der Fachbereichsleitung war sozusagen der Abschlussbericht eines Evaluations- und Entwicklungsprojekts. Währenddessen erfolgte beim hochschulinternen Verfahren die Dokumentation der Umsetzung der gutachterlichen Empfehlungen vornehmlich zeitnah und en passant im Rahmen des Protokolls der Auswertenden Veranstaltung, sozusagen als Zwischenstand von im Idealfall kontinuierlichen Veränderungsprozessen. Vergleichbare Umsetzungsberichte eineinhalb Jahre nach der Begutachtung fehlen.

Quellenangaben und Kontaktdaten

- Braun, J., Garve, M., & Reihlen, M. (2017): Qualitätsentwicklung im Netzwerk am Beispiel des Verbundprojekts Quality Audit. HOSL 3 61 17 07, C 3.14.
- Bornmann, L., Daniel H.-D., Mittag, S., & Mutz, R. (2004). Interne und externe Evaluation. Wichtige Voraussetzungen und praktische Handreichungen zur Gestaltung mehrstufiger Evaluationsverfahren. HOSL 1 00 04 06, E 7.1.
- Mittag, S. (2006). Qualitätssicherung an Hochschulen. Eine Untersuchung zu den Folgen der Evaluation von Studium und Lehre. Münster [u.a.]: Waxmann.
- Mittag, S.; Bornmann, L. & Daniel, H.-D. (2006). Qualitätssicherung und -verbesserung von Studium und Lehre durch Evaluation. Akzeptanz und Folgen mehrstufiger Evaluationsverfahren. Beiträge zur Hochschulforschung, Heft 2, 28. Jahrgang, 2006, 6-27.
- Universität Greifswald (Hrsg. 2012): Regelmäßige interne und externe Evaluation der Lehreinheiten an der Universität Greifswald gem. § 3a LHG M-V („Periodische Fachevaluation“) - Prozessbeschreibung und Programmablaufplan. Universität Greifswald: Integrierte Qualitätssicherung Studium und Lehre. Online verfügbar [09.09.2020]: https://www.uni-greifswald.de/storages/uni-greifswald2_Studium/2.1_Studienangebot/2.1.4_Qualitaet_in_Studium_und_Lehre/Prozessbeschreibung_Qualitaetshandbuch_20200107_U-Greifswald_Prozessbeschreibung_externer_Fachevaluation.pdf

Autor*innen:
Dr. Andreas Fritsch, Jessica Walter, Lara Lichtenthäler
Universität Greifswald
Integrierte Qualitätssicherung in Studium und Lehre
Baderstraße 4/5 - D-17489 Greifswald

Wir sind offen für Fragen und Anregungen. Bei Interesse stellen wir Ihnen gern weitere Ergebnisse zur Verfügung.
Internet: www.uni-greifswald.de/iqs,
E-Mail: qualitaetssicherung@uni-greifswald.de

Ergebnisse

Insgesamt wurden 238 von 251 erfassten gutachterlichen Empfehlungen weiterverarbeitet. Die meisten gutachterlichen Empfehlungen bezogen sich auf die Ausstattung ($\approx 19\%$) des jeweiligen Fachbereichs sowie auf die Positionierung und Profilbildung ($\approx 15\%$). Ca. 13 % der Empfehlungen betrafen die Planung und Organisation von Studium und Lehre. Die wenigsten Empfehlungen thematisierten die Lehr- und Lernformen ($\approx 5\%$), die Förderung wissenschaftlichen Nachwuchses ($\approx 5\%$) sowie Bildungs- und Ausbildungsziele ($\approx 2\%$) (Abbildung 1).

Prozentual gesehen wurden die meisten Empfehlungen in den Themenbereichen Bildungs- und Ausbildungsziele, Lehr- und Lernformen sowie Prüfungen umgesetzt bzw. hierfür Maßnahmen ergriffen (jeweils $> 50\%$). In den Bereichen Ausstattung (50 %), Wissenschaftlicher Nachwuchs (45 %) Verwaltung und akademische Selbstverwaltung (je 38 %) wurden relativ viele Empfehlungen begründet nicht umgesetzt. In den Themengebieten Qualitätssicherung (47 %), Positionierung und Profilbildung (44 %) sowie Lehrinhalte (38 %) wurden relativ häufig keine Maßnahmen ergriffen, um die gutachterlichen Empfehlungen umzusetzen.

Es zeigt sich, dass in den hochschulintern organisierten Verfahren 40 % der gutachterlichen Empfehlungen umgesetzt worden sind oder aber Maßnahmen zur Umsetzung eingeleitet wurden. Bei dem im Verbund organisierten mehrstufigen Evaluationsverfahren wurde eine höhere Rate (56 %) der umgesetzten gutachterlichen Empfehlungen dokumentiert. Der Anteil der gutachterlichen Empfehlungen, die begründet nicht umgesetzt wurden, beträgt bei den intern organisierten Verfahren rund 29 %, bei den Verbundevaluationsverfahren nur 15 %. Der Anteil der gutachterlichen Empfehlungen, für deren Umsetzung keine Maßnahmen ergriffen wurden, fällt mit 32 % und 29 % in beiden Evaluationsverfahren ähnlich hoch aus (Abbildung 2).

Abbildung 1: Der Bearbeitungsgrad der gutachterlichen Empfehlungen aus 25 mehrstufigen Evaluationsverfahren zwischen 2012 und 2017 nach Themenbereichen

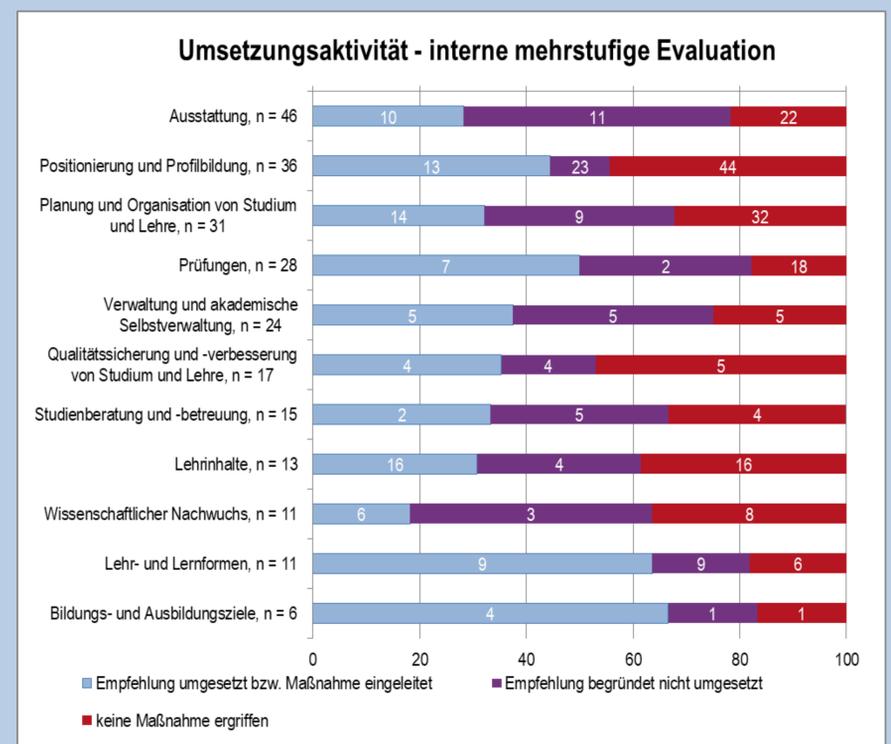
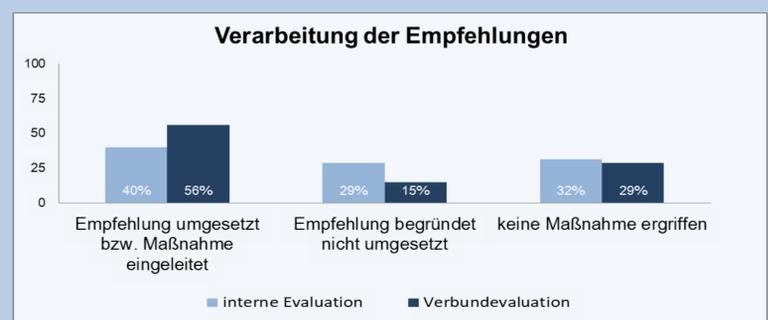


Abbildung 2: Bearbeitungsgrad der gutachterlichen Empfehlungen – Vergleich des hochschulinternen mehrstufigen Evaluationsverfahren mit dem im Verbund organisierten Verfahren



Handlungsempfehlungen

Mit den Ergebnissen lässt sich nicht belegen, dass die Wirksamkeit eines der beiden Evaluationsverfahren höher oder niedriger ist. Entscheidend für die Beurteilung der Wirksamkeit erscheint dabei jedoch die Kategorie „Keine Maßnahme ergriffen“. Die Kategorien „Empfehlungen umgesetzt bzw. Maßnahme eingeleitet“ und „Empfehlung begründet nicht umgesetzt“ verhalten sich komplementär zueinander.

Die verwendeten Methoden der Datenerhebung haben sich insgesamt als objektiv und reliabel erwiesen. Die von Mittag (2006) getroffenen Zuordnungsregeln lassen sich verlässlich anwenden. In Bezug auf die angezielte vergleichende Beurteilung der Wirksamkeit mehrstufiger Evaluationsverfahren sollten die Dokumentationen zur Umsetzung gutachterlicher Empfehlungen vergleichbaren Standards folgen.

Die mit einem Evaluationsverfahren im Hochschulverbund verbundenen erweiterten Vergleichs- und Lernmöglichkeiten sollten in geeigneter Weise auch in hochschulintern durchgeführten mehrstufigen Evaluationsverfahren integriert werden. Während die interne und externe Begutachtung vor allem auf die Bewertung von Qualitätskriterien und das Ableiten gutachterlicher Empfehlungen abzielt, erhöht die Einbeziehung von Hochschulverbänden und -netzwerken wie das Netzwerk Quality Audit (Braun/Garve/Reihlen, 2017) vor allem in der Follow-up-Phase der mehrstufigen Evaluationsverfahren das vorhandene Wissen und die Erfahrungen zur Umsetzung der gutachterlichen Empfehlungen. Das Benchmarking mit oder das Modelllernen an vergleichbaren Fachbereichen anderer Hochschulen fördert die Umsetzungscompetenz zu den gutachterlichen Empfehlungen.