

Open Access Evaluation:  
Lehr-Evaluation-Online (LEO) als Instrument zur studentischen Lehrveranstaltungsevaluation

Stefan Janke<sup>1</sup>, Julia Hein<sup>1</sup>, Martin Daumiller<sup>2</sup>, Raven Rinas<sup>2</sup>,  
Edgar Erdfelder<sup>1</sup>, Markus Dresel<sup>2</sup> & Oliver Dickhäuser<sup>1</sup>  
<sup>1</sup> Universität Mannheim; <sup>2</sup> Universität Augsburg

**Kontaktadressen – Autor\*innen**

1. Stefan Janke  
Dr., Wissenschaftlicher Mitarbeiter,  
Universität Mannheim,  
LS für Pädagogische Psychologie  
A 5, 6 | Raum B 313 | 68159 Mannheim  
Mail: stefan.janke@uni-mannheim.de

3. Martin Daumiller  
Dr., Akademischer Rat aZ,  
Universität Augsburg, LS für Psychologie  
Universitätsstr. 10 | Raum D 2087 |  
86135 Augsburg  
Mail: martin.daumiller@phil.uni-augsburg.de

5. Edgar Erdfelder  
Prof. Dr., Lehrstuhlinhaber,  
Prorektor für Forschung und  
Wissenschaftlichen Nachwuchs  
Universität Mannheim, LS für Kognitive und  
Differentielle Psychologie  
Schloss | Raum EO 255 | 68161 Mannheim  
Mail: erdfelder@psychologie.uni-  
mannheim.de

7. Oliver Dickhäuser  
Prof. Dr., Lehrstuhlinhaber  
Universität Mannheim,  
LS für Pädagogische Psychologie,  
A 5, 6 | Raum B 312 | 68159 Mannheim  
Mail: oliver.dickhaeuser@uni-mannheim.de

2. Julia Hein  
M.Sc., Wissenschaftliche Mitarbeiterin,  
Universität Mannheim,  
LS für Pädagogische Psychologie  
A 5, 6 | Raum B 314 | 68159 Mannheim  
Mail: julia.hein@uni-mannheim.de

4. Raven Rinas  
M.Sc., Wissenschaftliche Mitarbeiterin,  
Universität Augsburg, LS für Psychologie  
Universitätsstr. 10 | Raum D 2072 |  
86135 Augsburg  
Mail: raven.rinas@phil.uni-augsburg.de

6. Markus Dresel  
Prof. Dr., Lehrstuhlinhaber  
Universität Augsburg, LS für Psychologie,  
Universitätsstr. 10 | Raum D 2086 |  
86135 Augsburg  
Mail: markus.dresel@phil.uni-augsburg.de

**Entwurfssfassung:** Dezember, 2020. Dieser Entwurf kann teilweise von der Finalfassung des im Webler Verlag publizierten Artikel abweichen. Er ist keine Kopie des publizierten Artikels.

**Abstract:** We present an open-access platform, Lehr-Evaluation-Online (LEO; Teaching Evaluation Online), which provides higher education faculty with an option to conduct student teaching evaluations free of charge. LEO is based on a German adaption of the “Student Evaluation of Educational Quality” (Marsh, 1982; 2007). First data collected with LEO during a testing phase supports that the test scores can be interpreted in a valid way. The online tool presents data in a user-friendly way that enables instructors to evaluate their own teaching based on criterial and social comparisons. We consider LEO to be a useful tool that supports higher education faculty in maintaining scientific standards of teaching evaluations.

**Studentische Lehrveranstaltungsevaluationen werden an vielen Hochschulstandorten mit Hilfe von Online-Befragungsinstrumenten erfasst und verarbeitet (Treischl & Wolbring, 2017). Dies ermöglicht eine schnelle Rückmeldung an die Dozierenden. Entsprechende Plattformen müssen allerdings auch gepflegt und Vergleichsdaten auf dem neuesten Stand gehalten werden, um sinnvolle Rückmeldungen liefern zu können. Wir stellen in diesem Artikel ein Befragungsinstrument vor, das es Hochschulstandorten ermöglicht, diese Aufgaben auszulagern. Lehr-Evaluation-Online (kurz: LEO) ist ein offen online zugängliches Tool (Open Access) zur Erfassung und Rückmeldung studentischer Veranstaltungsbewertungen. Es basiert auf einem international validierten Fragebogen zur studentischen Lehrveranstaltungsevaluation, ist adaptierbar und liefert den Dozierenden Rückmeldung zu multiplen Bewertungsdimensionen, auch unter der Nutzung von Vergleichsdaten. LEO entspricht damit dem aktuellen Stand der Forschung zu studentischer Veranstaltungsbewertung und steht den nutzenden Hochschulen kostenfrei zur Verfügung.**

### **1 Was zeichnet LEO als onlinebasierten Fragebogen aus?**

LEO ist ein Instrument zur Durchführung und Rückmeldung studentischer Lehrveranstaltungsevaluationen. Studentische Lehrveranstaltungsevaluationen sind anonyme Befragungen von Studierenden zu Qualitätsmerkmalen der von ihnen besuchten Veranstaltung (Marsh, 2007). Die Ergebnisse solcher Befragungen sollen (über unmittelbare Rückmeldungen der Studierenden in der Lehrveranstaltung hinaus) Dozierenden die Sicht ihrer Studierenden auf die jeweilige Veranstaltung vor Augen führen. Auf Basis dieser Rückmeldung können Dozierende ihre Lehre und das Geschehen in den Veranstaltungen reflektieren und eigene Entwicklungsbereiche identifizieren. Ergebnisse studentischer Lehrveranstaltungsevaluationen werden darüber hinaus auch in der hochschuleigenen Personalauswahl eingesetzt (Linse, 2017; Treischl & Wolbring, 2017). Die hinter dieser Verwendung stehende Logik, dass studentische Lehrveranstaltungsevaluationen tatsächlich geeignet sind, die Güte von Lehre zu indizieren, wird in der Literatur breit diskutiert (siehe u.a. Spooren et al., 2013). Es ist jedoch nicht von der Hand zu weisen, dass studentische Lehrveranstaltungsevaluationen weniger von selbstwertdienlichen Attributionsfehlern verzerrt sind als rein selbstreflektive Verfahren (Spinath & Brünken, 2016). Entsprechend stellen sie ein (mit Vorsicht zu betrachtendes) Mittel dar, um mögliche Probleme in der eigenen Lehre aufzudecken und Rückmeldungen über verschiedene Lehrkonzepte zu erhalten. Auf dieser Basis ist es möglich, die richtigen Maßnahmen, wie beispielsweise hochschuldidaktische Schulungen, zu ergreifen (Blair & Valdez Noel, 2014). Gleichzeitig sollten bei der Durchführung studentischer Lehrveranstaltungsevaluationen allerdings auch deren Limitationen berücksichtigt werden. Dies kann beispielsweise durch Miterhebung bekannter Biasvariablen, wie beispielsweise dem Besuchsgrund (Wahl- oder Pflichtkurs), dem Ausmaß von Vorwissen und Interesse für den Gegenstand geschehen (Marsh & Roche, 1997; Rindermann, 2003; Spooren, Brockx & Mortelmans, 2013).

Obwohl studentische Lehrveranstaltungsevaluationen wichtige Werkzeuge für die Weiterentwicklung von Hochschullehre darstellen können, gibt es limitierende Faktoren in der Verfügbarkeit von Instrumenten und Verfahren zur Lehrveranstaltungsevaluation für einzelne Dozierende. So vertritt etwa Bialowas (2016) die Position, dass die systematische Durchführung von Lehrveranstaltungsevaluationen einer umfassenden Betreuung durch ein an den Hochschulen institutionalisiertes Qualitätsmanagement bedarf. Es gibt jedoch beträchtliche Unterschiede in den Ressourcen, die Hochschulen und Universitäten für die Ausstattung solcher Stabsstellen für Qualitätsmanagement zur Verfügung stellen (können). Weitergehend ist es auch von hoher Bedeutung, dass genutzte Evaluationssoftware gewartet und die

verwendeten Inventare stets entsprechend neuester wissenschaftlicher Erkenntnisse angepasst werden. Hierzu ist es notwendig, wissenschaftliche Expert\*innen dauerhaft in den Evaluationsprozess einzubeziehen (siehe Rindermann, 2003; aber auch Janke & Mühlbauer, 2019). Das notwendige wissenschaftliche Personal mit Expertise im Bereich Lehrevaluationsforschung und/oder Hochschuldidaktikforschung steht dabei wahrscheinlich noch weniger Standorten zur Verfügung als die finanziellen Mittel zum Aufbau eines tragkräftigen Qualitätsmanagements.

Diese Ressourcenbeschränkung auf Seiten einzelner Hochschulen lässt sich überwinden: Wir zielen mit LEO darauf ab, allen Hochschulen und Hochschuldozierenden (im deutschsprachigen Raum) unabhängig von den jeweiligen örtlichen Gegebenheiten ein Evaluationsinstrument zugänglich zu machen, das wissenschaftlichen Gütekriterien entspricht. Die Zugänglichkeit zu diesem Instrument erfolgt online und kostenfrei (Open Access). Als inhaltliche Grundlage dient die Adaption eines ursprünglich englischsprachigen Fragebogens mit dem Titel „Student Evaluation of Educational Quality“ (SEEQ; Marsh, 1982; 2007). Es handelt sich dabei um das im englischsprachigen Raum am verbreitetste und am besten validierte Instrument zur studentischen Lehrveranstaltungsevaluation.

### 1.1 Multidimensionalität von LEO: der SEEQ als inhaltliche Grundlage

Der SEEQ basiert auf einer multidimensionalen Auffassung von Lehrqualität (Marsh & Roche, 1997; siehe auch Rindermann, 2016). Dies bedeutet, dass Lehrqualität nicht als unidimensionales Konstrukt (zum Beispiel gemessen als studentische Leistung in Tests) aufgefasst wird, sondern mehrere bedeutsame Merkmalsdimensionen erfasst werden. Die im Rahmen von LEO eingesetzte Adaption des SEEQ erfasst Lehrqualität auf Prozess- und Produktebene. Auf der Produktebene wird Lehrqualität indiziert durch den selbst eingeschätzten *Lernzuwachs der Studierenden*, die *Gesamtbeurteilung der Lehrveranstaltung kriterial* (Einzelitem, als Schulnote von 1-5), die *Gesamtbeurteilung der Lehrperson kriterial* (Einzelitem, als Schulnote von 1-5), die *Gesamtbeurteilung der Lehrperson normativ im Vergleich zu anderen Lehrpersonen* (Einzelitem, als Schulnote von 1-5) und die *Gesamtbeurteilung der Lehrveranstaltung normativ im Vergleich zu anderen Lehrveranstaltungen* (Einzelitem, als Schulnote von 1-5). Auf der Prozessebene werden Qualitätsmerkmale der Instruktion erfasst. Dies sind das *Engagement der Lehrperson*, die *Stoffstrukturierung und -präsentation*, die *Aktivierung der Studierenden*, das *Sozialklima*, die *Stoffbreite*, die *Qualität der Leistungsbewertung* und die *Aufgabenqualität*.

Neben diesen Variablen auf Dozierendenseite werden auch Biasvariablen in Bezug auf Rahmenbedingungen der Veranstaltung und Eigenschaften der teilnehmenden Studierenden erfasst, welche die Evaluationsergebnisse positiv sowie negativ beeinflussen und entsprechend bei der Interpretation der Ergebnisse berücksichtigt werden sollten (Marsh & Roche, 1997; Spooren, Brockx & Mortelmans, 2013). Auf Seite der Studierenden erfassen je eine Subskala die *Qualität studentischer Beiträge* sowie Einzelitems die *Leistungsaspirationen*, das *Interesse*, das *Vorwissen* und den *Besuchsgrund* (Pflicht, Wahlpflicht, freiwillig). In Bezug auf die Rahmenbedingungen wird mit einer Subskala das *Anspruchsniveau* erfasst, sowie mit einem Einzelitem die für die Veranstaltung aufgewendeten *Arbeitsstunden pro Woche*.

Wenn nicht anders beschrieben, erfolgt die Erfassung der einzelnen Indikatoren von Lehrqualität mit Hilfe von Subskalen, die Items mit fünfstufigen Likert-Antwortskalen umfassen. Diese Skalen haben sich international bewährt und wurden auch in Deutschland erfolgreich erprobt (dokumentiert bei Dresel, Engelschalk, Grassinger, 2012 und Daumiller, Grassinger, Engelschalk & Dresel, 2020). Der SEEQ wird auf der LEO-Plattform durch weitere offene Fragen ergänzt. Dabei handelt es sich um die Fragen „Was hat Ihnen an der Lehrveranstaltung besonders gut gefallen?“ und „Was ließe sich Ihrer Meinung nach an der Lehrveranstaltung noch verbessern?“. Außerdem umfasst das Tool ein weiteres offenes Feld

für *“Sonstige Bemerkungen zu dieser Lehrveranstaltung”*. Studentische Angaben zu diesen offenen Fragen sollen die quantitative Rückmeldung mit qualitativen Verbesserungsvorschlägen ergänzen. Weitergehend haben Nutzer\*innen von LEO die Möglichkeit den bestehenden Fragebogen um eigene geschlossene oder offene Fragen zu ergänzen. Um den eingesetzten Fragebogen Hochschuldozierenden offen zugänglich zu machen, setzt LEO auf eine Open Access Plattform.

## 1.2 Technische Umsetzung von LEO: Die Open Access Plattform

Dozierende können sich unter <https://www.lehr-evaluation-online.de/> anmelden, um selbstständig Veranstaltungen zur Evaluation anzulegen<sup>1</sup>. Nach Anmeldung erhalten Dozierende auf der Startseite einen Überblick, über die bereits angelegten und evaluierten Veranstaltungen (siehe Abb. 1). Es stehen abhängig vom Fortschritt der Evaluation verschiedene Funktionen zur Verfügung (neue Veranstaltung anlegen, angelegte Veranstaltungen bearbeiten, Ergebnisse bereits evaluierter Veranstaltungen anzeigen oder herunterladen, Rohdaten bereits evaluierter Veranstaltungen exportieren).

[Abbildung 1 etwa hier platzieren]

Die Veranstaltungen werden mit ihrem Veranstaltungsnamen zur Evaluation hinterlegt; weiterhin werden Angaben zur Art der Veranstaltung und dem Fach abgefragt. Diese Angaben werden es zukünftig ermöglichen, für eine Vielzahl von Fächern und alle Veranstaltungsarten getrennte Normdaten in der Ergebnisdarstellung anbieten zu können; aktuell ist diese Funktion noch nicht vollständig implementiert. Der Zeitpunkt, zu welchem im Semester die Evaluation durchgeführt werden soll, sowie die Länge des Evaluationszeitraums, kann frei gewählt werden. Als Sprache der Fragebögen für die Studierenden steht neben Deutsch auch Englisch zur Verfügung, damit auch englischsprachige Veranstaltungen evaluiert werden können. Außerdem ermöglicht LEO teilweise individualisierte Lehrveranstaltungsevaluationen abhängig vom Erkenntnisinteresse der Dozierenden. Dies bedeutet, dass Dozierende eigene offene und geschlossene Fragen zur Lehrveranstaltungsevaluation hinzufügen können. In diesem Rahmen sind entsprechend Anpassungen an die Spezifika verschiedener Hochschulorte und Fächer möglich. Darüber hinaus können zur Evaluation digitaler Lehrformate vorkonstruierte geschlossene Fragen wahlweise ergänzt werden. Hierbei werden die Nützlichkeit der eingesetzten digitalen Lehrformate, technischen Rahmenbedingungen und allgemeine Zusatzinformationen zu digitalen Lehrformaten in Einzelitems erfragt. Ausführliche Information zu dem eingesetzten Fragebogen und den möglichen Ergänzungen sind online verfügbar<sup>2</sup>. Die durchschnittliche Bearbeitungszeit des Fragebogens durch die Studierenden beträgt 5 bis 8 Minuten und verlängert sich durch zusätzliche Fragen entsprechend.

Die studentische Lehrveranstaltungsevaluation mit LEO ist sowohl am PC als auch mittels mobiler Endgeräte möglich. Die Studierenden benötigen einen individuellen Zugangscode, um auf der Plattform ihre Veranstaltung zu evaluieren. So wird sichergestellt, dass Studierende nicht mehrfach an der Evaluation teilnehmen und nur Studierende die Veranstaltung bewerten, die diese auch besuchen. Die Dozierenden können sich zwischen zwei Optionen entscheiden, wie die Studierenden zur Teilnahme an der Evaluation eingeladen

---

<sup>1</sup> LEO wurde von Jan Siebert in der Programmiersprache MySQL programmiert und von Hannes Kohlsaat weiterentwickelt. Ausgangspunkt war der Quellcode der ebenfalls kostenfreien und mit ähnlicher Funktionalität ausgestatteten *Internetbasierten studentischen Evaluation von Lehrveranstaltungen* (InstEvaL), die u.a. an der Universität Mannheim von 2003 bis 2018 Verwendung fand. InstEvaL wurde von Edgar Erdfelder, Christoph Stahl und Stephanie Pesch entwickelt und von Volker Gropp an der Universität Mannheim implementiert.

<sup>2</sup> [https://www.lehr-evaluation-online.de/static/Information\\_zum\\_Fragebogen.pdf](https://www.lehr-evaluation-online.de/static/Information_zum_Fragebogen.pdf)

werden sollen. Die erste Möglichkeit besteht darin, die Teilnehmer\*innen automatisiert zu einem bestimmten Zeitpunkt per E-Mail einzuladen. In diesem Fall erhalten die Studierenden ihren Code in einer E-Mail. Um dies zu ermöglichen, ist es erforderlich, dass die Mailadressen der Studierenden in der Evaluationsplattform hinterlegt werden. Die E-Mail-Adressen werden separat von den Evaluationsdaten zwischengespeichert, ausschließlich zum Zwecke der Einladung zur Lehrveranstaltung genutzt und sofort nach Abschluss des Evaluationszeitraums automatisch gelöscht. Studierende werden automatisiert erinnert, wenn sie in einem bestimmten Zeitraum noch nicht an der Lehrveranstaltungsevaluation teilgenommen haben. Die zweite Möglichkeit ist eine Einladung mit Hilfe auf Papier gedruckter individueller Codes, welche Dozierende in ihrer Veranstaltung verteilen können. Hierzu müssen die Dozierenden im Vorfeld im System die Anzahl der Studierenden angeben, woraufhin die notwendige Anzahl Codes automatisch generiert und zur Verfügung gestellt wird. Zur Erhöhung der Teilnahmequote ist es denkbar, den Studierenden am Ende der eigenen Lehrveranstaltung vor Ort Zeit zum Ausfüllen des Evaluationsfragebogens auf dem Smartphone oder am Laptop zu geben.

Die über LEO erhobenen Daten werden auf Servern der Universitäten Mannheim, Ulm, Freiburg und Karlsruhe und somit innerhalb Deutschlands in der gemeinsamen bw-cloud gespeichert. Die Daten werden zum Schutz gegenüber Dritten mit HTTPS-Technologie und einem 2048-Bit Schlüssel den aktuellen Sicherheits-Empfehlungen entsprechend gesichert und nicht an Dritte weitergegeben. Die aggregierten Daten werden nur in anonymer Form zur weiteren Verbesserung der LEO Plattform und zu Forschungszwecken genutzt. Ein Rückschluss auf einzelne Personen oder Veranstaltungen erfolgt hierbei nicht.<sup>3</sup>

### 1.3 Hochschuldidaktische Elemente von LEO: Die Ergebnissrückmeldung

Nach Ende des Evaluationszeitraumes werden die Dozierenden darüber benachrichtigt, dass sie die automatisch erstellten Ergebnissrückmeldungen online einsehen können. Die Dozierenden erhalten die Ergebnisse außerdem auch via Mail als PDF zugesendet<sup>4</sup>. Die Ergebnissrückmeldung enthält leicht interpretierbare Rückmeldungen zu den Skalenwerten von LEO und Verteilungen der Antwortmuster einzelner Items. Darüber hinaus ist ein Vergleich der eigenen Lehrqualität mit Normdaten aus 3660 Veranstaltungen in bildungs-, geistes-, natur-, sozial- und sprachwissenschaftlichen Studienfächern möglich (siehe Abb. 2), welche mit Hilfe der Papiervariante des SEEQ evaluiert wurden (n = 76687 Studierende). Diese Normdaten ermöglichen zum jetzigen Zeitpunkt eine Einordnung in eine große Gesamtheit von Lehrveranstaltungen, allerdings noch keine fachspezifische Einordnung.

Bei der Entwicklung von LEO wurde darauf geachtet, motivationsförderliche Elemente in die Rückmeldung zu integrieren. Diese Elemente sollen Dozierende anregen über die Evaluation und mögliche Folgemaßnahmen nachzudenken, da allein der Erhalt einer Evaluationsrückmeldung die Lehrqualität nicht automatisch verbessert (u.a. Bialowas, 2016). So werden auf der Seite zur Ergebnissrückmeldung allgemeine Informationen zur Verbesserung der Lehrqualität als niederschwelliges Informationsangebot angeboten (zur Bedeutung von Beratung im Kontext von Lehrveranstaltungsevaluationen siehe Dresel & Rindermann, 2011). Außerdem wird bei der Ergebnissrückmeldung der einzelnen Items (Häufigkeitsdiagramm) auf ein Farbschema mit Ampelfarben zurückgegriffen, um die inhaltliche Verarbeitung des quantitativen Feedbacks zu erleichtern. So wird eine eher positive Bewertung (4-5 auf der fünfstufigen Skala) mit grünen Häufigkeitsbalken, eine befriedigende Bewertung (3 auf der fünfstufigen Skala) mit gelben Häufigkeitsbalken und eine eher negative Bewertung (1-2 auf

---

<sup>3</sup> Die vollständige Datenschutzerklärung ist online verfügbar (<https://www.lehr-evaluation-online.de/legal/privacy>).

<sup>4</sup> Eine Musterrückmeldung kann online eingesehen werden unter: <https://www.lehr-evaluation-online.de/static/Musterrergebnisse-LEO.pdf>

der fünfstufigen Skala) mit roten Häufigkeitsbalken indiziert. Dies soll es den Dozierenden erleichtern, auf einen Blick die inhaltliche Bedeutung des studentischen Antwortmusters zu erfassen und mögliche Ansatzpunkte zur Verbesserung der Lehrveranstaltung schneller und leichter zu identifizieren.

[Abbildung 2 etwa hier platzieren]

## 2 Erprobungsphase: LEO im Feld

Seit 2019 ist LEO offen für Dozierende verschiedener Hochschulen zugänglich. LEO lässt sich bedingungs- und kostenfrei nutzen. Die ersten Daten zur Nutzung von LEO sind vielversprechend. Zur Vermeidung von Verzerrungen durch die abrupte Umstellung auf digitale Hochschullehre im Sommersemester 2020 ziehen wir für die folgenden Ausführungen zur Güte des Instrumentes die Datengrundlage vor dem Eintritt der COVID-19 Pandemie heran. Anfang März 2020 gab es 796 registrierte Nutzer\*innen. Erste Hintergrunddaten zur Nutzung des Evaluationsinstrumentes lagen für 2939 Studierenden in 279 Kursen von 185 Dozierenden vor. Im Durchschnitt nahmen 37.3% der eingeladenen Studierenden an der jeweiligen Evaluationsbefragung teil. Zur Beurteilung der Teilnahmequote ist zu bedenken, dass LEO in der Erprobungsphase häufig als zusätzliche Evaluation – neben weiteren Evaluationsmaßnahmen – durchgeführt wurde. Eine finale Abschätzung der mit LEO zu erzielenden Teilnahmequote ist deshalb zum jetzigen Zeitpunkt noch nicht möglich.

Zur Prüfung der Reliabilität des eingesetzten Instrumentes konnten wir den anonymisierten Gesamtdatensatz nutzen. Auf Basis dieses Datensatzes überprüften wir, ob die Subskalen des SEEQ-DE – ähnlich wie bereits in vorangegangenen Untersuchungen gezeigt (Daumiller et al., 2020) – auch bei ihrem Einsatz in LEO eine hinreichende Messgenauigkeit erreichen. Dies war der Fall ( $.90 \geq \text{Cronbachs } \alpha \geq .83$ ). Darüber hinaus überprüften wir auch die Abgrenzbarkeit der einzelnen Skalen mit Hilfe einer latenten konfirmatorischen Faktorenanalyse. Dabei ließ sich eine akzeptable Passung der angenommenen Faktorstruktur an die Daten aufzeigen; Modelgüte:  $\chi^2(518) = 2363.96$ ,  $p < .001$ , RMSEA = .04, SRMR = .04; CFI = .95, TLI = .95 (zur Interpretation siehe Hu & Bentler, 1999) und die Items zeigten zufriedenstellende Faktorladungen (mittleres  $\lambda = .76$ ). Aus diesem Befundmuster schließen wir, dass sich der SEEQ-DE als inhaltliche Grundlage der LEO Plattform durch hohe Reliabilität und Konstruktvalidität auszeichnet. In einer kurzen Befragung der Dozierenden nach der ersten Evaluation mit LEO schätzten diese ihre Zufriedenheit mit den Ergebnissen der Evaluation ( $M = 6.42$ ,  $SD = 1.63$ , Skalenrange: 1-8) und darüber, was sie aus der Evaluation lernen konnten ( $M = 6.43$ ,  $SD = 1.68$ , Skalenrange: 1-8) auf Einzelitems im Durchschnitt als hoch ein. Außerdem berichten 67.3% der teilnehmenden Dozierenden, mindestens einen konkreten Verbesserungsvorschlag für ihre zukünftige Lehre aus der Evaluation mittels LEO ziehen zu können. Daraus schließen wir, dass LEO zumindest aus Sicht der Anwender\*innen eine Grundlage für Veränderungsprozesse mit dem Ziel der Verbesserung der Lehrqualität schaffen kann.

## 3 Schlussbetrachtungen: LEO als Instrument des Qualitätsmanagements in der Hochschullehre

LEO richtet sich ganz gezielt an Hochschulen und Dozierende, die auf der Suche nach einem erprobten und gut validierten Instrument zur Lehrveranstaltungsevaluation sind. LEO bietet diesen Institutionen und Personengruppen einen effizienten und kostenfreien Zugang zu einem international verbreiteten und auch in der deutschen Fassung messgenauen und gültigen Instrument zur Evaluation der eigenen Hochschullehre. Die Oberfläche ist benutzerfreundlich angelegt und die generierten Rückmeldungen ermöglichen den kriterialen und normorientierten

Vergleich. LEO kann eine Steigerung der Qualität der Lehre durch Reflektion studentischen Feedbacks auch dort ermöglichen, wo ein Mangel an finanziellen oder personellen Ressourcen die Nutzung von Lehrevaluationen normalerweise erschwert. Auch für Institutionen mit bereits eigenen (u.U. bislang nicht validierten und mitunter kostspieligen) Verfahren ist LEO eine sinnvolle Alternative.

LEO bietet sowohl die Möglichkeit, die Studierenden mit Hilfe von Einladungsmails als auch mittels persönlicher Codes zu der Evaluation einzuladen. Die letztere Option ermöglicht es Dozierenden, die Plattform auch dann zu nutzen, wenn ihnen die Mailadressen der Studierenden nicht zur Verfügung stehen oder die Studierenden keine Einwilligung zur Verwendung ihrer Mailadressen zwecks Eingabe in LEO geben. An dieser Stelle soll darauf hingewiesen werden, dass es in der Verantwortung der LEO nutzenden Dozierenden liegt, zu überprüfen, ob vorhandene Evaluationssatzungen eine Nutzung der hochschuleigenen Mailadresse der Studierenden für Zwecke der Lehrevaluation erlauben oder die Einholung entsprechender Einverständniserklärungen notwendig ist. Auch die Sicherstellung optimaler Teilnahmeraten zur Vermeidung von Selektionseffekten liegt in der Verantwortung der Dozierenden. LEO kann in dieser Hinsicht allerdings durch Erinnerungsmails unterstützen, falls Mailanschriften als Kontaktform gewählt wurden. Darüber hinaus erhalten alle LEO nutzenden Dozierenden vor Durchführung der Evaluation den Hinweis via Mail, dass die Präsenzbearbeitung der Online-Evaluation in einer der letzten Sitzungen die Teilnahmequote erhöhen kann.

Es versteht sich von selbst, dass die flächendeckende Anwendung von LEO allein kein Garant zur Verbesserung von Hochschullehre sein kann. Hierzu bedarf es weitergehender hochschuldidaktischer Weiterbildungs- und Beratungsangebote. Diese Angebote sollten entweder vor Ort an den jeweiligen Hochschulen geschaffen werden oder Dozierenden zielgerichtet über Hochschulverbünde zugänglich gemacht werden. Für Dozierende die (noch) keinen Zugang zu solchen Angeboten haben, bietet LEO jedoch angebunden an die Ergebnismeldung allgemeine niederschwellige Informationen zu digitalen Medien sowie Büchern zu hochschuldidaktischen Maßnahmen. Darüber hinaus ermöglicht die einfache Zugänglichkeit der LEO-Ergebnisse (Ampel-Symbolik) den Nutzer\*innen optimale Weiterbildungsangebote zu identifizieren und eine Gesprächsgrundlage für Beratungen herzustellen.

LEO kann des Weiteren die grundsätzliche Kritik an der Verwendung studentischer Lehrevaluationsurteile als Maße der Qualität von Lehrveranstaltungen nicht vollends ausräumen. Ein wesentliches Ziel ist es jedoch den Informationsgehalt dieser Rückmeldung durch Einschluss bekannter studentischer Biasvariablen (Interesse am Gegenstand, Vorwissen, Grund des Besuchs) weiter zu erhöhen. So enthält LEO zum aktuellen Zeitpunkt zwar keine Option zur Bereinigung der studentischen Rückmeldung auf Basis dieser Biasvariablen, Informationen über Verteilungsmaße (Mittelwert, Standardabweichungen) ermöglichen es Nutzer\*innen aber das studentische Urteil auch im Lichte der Biasvariablen besser einzuordnen. Beispielsweise können Dozierende so besser abschätzen, ob sie es mit einem von Anfang an hochmotivierten Kurs zu tun hatten oder ob es ihnen gelungen ist, Lernfreude in Studierenden zu wecken, welche eigentlich kein Interesse an dem Lerngegenstand hatten und die Veranstaltung als reine Pflichtveranstaltung ansahen. Auch wenn eine weitere Rückkopplung der Evaluationsdaten an zentral gespeicherte Leistungsdaten zur (normbasierten) Einordnung von Vorwissen, eine noch objektivere Beurteilung der Evaluationsdaten ermöglichen würde, sind solche Maßnahmen datenschutzrechtlich wegen der Notwendigkeit der Deanonymisierung von Evaluationsdaten kritisch zu sehen. Entsprechend ist kein entsprechender Mechanismus für LEO angedacht.

Wir weisen darauf hin, dass zur optimalen Einordnung der Rückmeldungsdaten auch der Zeitpunkt der Datenerfassung (bspw. vor oder nach Erbringung einer Prüfungsleistung) berücksichtigt werden sollte. Da aus didaktischen Gründen je nach Veranstaltungskonzept

verschiedene Erhebungszeitpunkte sinnvoll sein können, möchten wir diesbezüglich keine Empfehlung abgeben. Wir empfehlen allerdings, dass Dozierende a) sich auf Basis ihres didaktischen Konzeptes Gedanken über den optimalen Messzeitpunkt machen und b) den Erhebungszeitpunkt durch die Auswahl des Evaluationszeitraumes homogenisieren (bspw. Teilnahme nur in der letzten Woche vor den Prüfungen möglich). LEO enthält zum aktuellen Zeitpunkt keine Möglichkeit um die erfassten Daten getrennt für verschiedene Messperioden auszugeben.

Als allgemein zugängliche und breit angelegte Plattform ist LEO nicht an die Besonderheiten einzelner Hochschulorte angepasst. Eine solche Anpassung kann zwar durch Hinzufügen einzelner geschlossener und offener Fragen durch die Nutzer\*innen erfolgen, wesentliche Subskalen des SEEQ lassen sich jedoch nicht deaktivieren. Wie bereits dargestellt, gibt es aus unserer Sicht allerdings auch gute wissenschaftliche Argumente für die Verwendung dieses Fragebogens im Rahmen studentischer Lehrevaluationen. Eine weitere Limitation in der Verwendung von LEO liegt darin, dass die Administration der einzelnen Befragungen nicht zentral über einen Account erfolgt, der beispielsweise von Mitarbeiter\*innen des hochschuleigenen Qualitätsmanagements verwaltet wird, sondern über die Zugänge einzelner Lehrpersonen. Die Lehrpersonen haben jedoch die Möglichkeit die Evaluationsergebnisse als PDF sowie die Rohdaten der eigenen Veranstaltungsevaluation mit Mitarbeiter\*innen des Qualitätsmanagements zu teilen. So kann das lokale Qualitätsmanagement auf Basis der gesammelten Daten eine eigene Datenbank zur Lehrqualität an der jeweiligen Hochschule erstellen, welche wiederum für aggregierte Analysen nutzbar ist. Ein solcher Schritt kann hilfreich sein, um hochschulübergreifende oder fachbereichsspezifische Entwicklungspotenziale zu identifizieren und gezielt hochschuldidaktische Weiterbildungsmaßnahmen zu schaffen. Da die flächendeckende Weiterleitung der Daten Akzeptanz auf Seiten der Hochschuldozierenden erfordert, ist es sicher sinnvoll diese im Sinne einer partizipativen Evaluation weiter in den Prozess einzubeziehen (vgl. Janke & Mühlbauer, 2019).

LEO soll auch zukünftig weiter optimiert werden. Als erster Schritt steht insbesondere die Sammlung umfassender Normdaten für verschiedene Fächer und Veranstaltungsformen an. Die Rückmeldungen der Dozierenden während der Pilotierungsphase im Jahr 2019 werden weitergehend zur Verbesserung der Plattform genutzt. Für die Zukunft hat LEO aus unserer Sicht das Potential, wissenschaftliche Standards für studentische Lehrveranstaltungsevaluationen flächendeckend sicherzustellen.

#### **4 Erklärung zu Interessen der Autor\*innen an der LEO Plattform**

Die Entwicklung von LEO wurde aus Mitteln der Deutschen Forschungsgemeinschaft im Rahmen eines von Markus Dresel und Oliver Dickhäuser eingeworbenen Kooperationsprojekts (Kennziffern DR 454/8-1 und DI 929/5-1) finanziert. LEO ermöglichte dabei die Beantwortung von Forschungsfragen zur Bedeutung von Dozierendenmotivation für die Interpretation von Lehrveranstaltungsevaluationen. Die weitere Bereitstellung von LEO an die Allgemeinheit erfolgt ohne dass zum aktuellen Zeitpunkt weitere Forschungsfragen mit Hilfe der Plattform verfolgt werden. Grundsätzlich ist allerdings angedacht die gesammelten Nutzungs- und aggregierten Evaluationsdaten auch in Zukunft für Forschungsfragen bezüglich der Rezeption von Lehrevaluationen durch Hochschuldozierende zu nutzen. Entsprechende Forschungsarbeiten sollen sowohl der Optimierung der Plattform als auch der Genese von bildungswissenschaftlichen Erkenntnissen dienen. Die Autor\*innen haben daneben keinerlei monetäres Interesse an der Plattform. So sind weder Nutzungsgebühren noch monetarisierende Werbeeinnahmen angedacht. Die Plattform ist nachhaltig auf Servern der Universitäten Mannheim, Ulm, Freiburg und Karlsruhe verankert und es ist zum derzeitigen Zeitpunkt nicht angedacht sie wieder vom Netz zu nehmen.

## 5 Literatur

- Bialowas N. (2016). Möglichkeiten und Grenzen der Evaluation der Lehre. In: Steirische Hochschulkonferenz (Hrsg.) *Qualität in Studium und Lehre* (S. 263-281). Springer VS: Wiesbaden
- Blair, E., & Valdez Noel, K. (2014). Improving higher education practice through student evaluation systems: Is the student voice being heard?. *Assessment & Evaluation in Higher Education*, 39, 879-894.
- Daumiller, M., Grassinger, R., Engelschalk, T., & Dresel, M. (2020). *SEEQ-DE: Konstruktion und Überprüfung einer deutschsprachigen Adaption des Instruments „Student Evaluation of Educational Quality“ (Marsh)*. Manuscript under review.
- Dresel, M., Engelschalk, T. & Grassinger, R. (2012). *Kurzbericht über die Konstruktion und Überprüfung einer deutschen Adaption des Fragebogens „Student Evaluation of Educational Quality“ (SEEQ) von Marsh (1982, 2007)*. Augsburg: Universität, Lehrstuhl für Psychologie.
- Dresel, M. & Rindermann, H. (2011). Counseling university instructors based on student evaluations of their teaching effectiveness: A multilevel test of its effectiveness under consideration of bias and unfairness variables. *Research in Higher Education*, 52, 717-737.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1-55.
- Janke, S. & Mühlbauer, M. (2019). Von Reaktanz zu Akzeptanz: Ein iterativer Ansatz zur Einbindung relevanter Expert\*innen- und Interessengruppen bei der Weiterentwicklung studentischer Lehrevaluationen. *Qualität in der Wissenschaft*, 13, 116-121.
- Linse, A. R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*, 54, 94-106.
- 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.
- Marsh, H.W., and Roche, L.A. (1997). Making students' evaluations of teaching effectiveness effective. *American Psychologist*, 52, 1187-1197.
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Dordrecht, Netherlands: Springer.
- Rindermann, H. (2003). Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. *Zeitschrift für Evaluation*, 2, 233-256.
- Rindermann, H. (2016): Lehrveranstaltungsevaluation an Hochschulen. In: Großmann, D. & Wolbring, T. (Hrsg.), *Evaluation von Studium und Lehre. Grundlagen, Methodische Herausforderungen und Lösungsansätze* (S. 227-262). Wiesbaden: Springer VS.
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching. The state of the art. *Review of Educational Research*, 83, 598-642.
- Treischl, E., & Wolbring, T. (2017). The Causal Effect of Survey Mode on Students' Evaluations of Teaching: Empirical Evidence from Three Field Experiments. *Research in Higher Education*, 58, 904-921.
- Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22-42.

Meine Veranstaltungen
Meine Daten
Logout

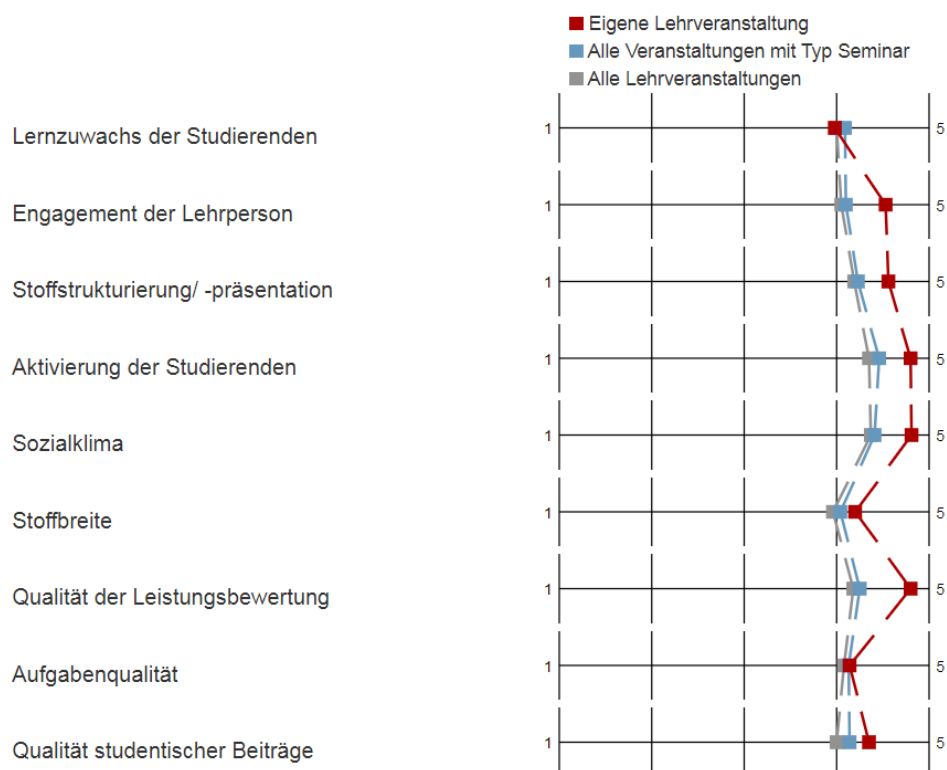
+ Neue Veranstaltung anlegen

### Meine Veranstaltungen

| Name der Veranstaltung                                                         | Teilnehmer(innen)                  | Status                                                      | Aktionen                                                                                                                                              |
|--------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| S1: Prävention und Intervention (Wintersemester 2020)                          | 5 Einladungen<br>0 Rückmeldungen   | Noch nicht begonnen<br>1. Dezember 2020 - 14. Dezember 2020 | <a href="#">Bearbeiten</a><br><a href="#">Teilnehmer(innen) hinzufügen</a><br><a href="#">Erzeugte Codes herunterladen</a><br><a href="#">Löschen</a> |
| S1: Ausgewählte Probleme der Pädagogischen Psychologie I (Sommersemester 2020) | 13 Einladungen<br>11 Rückmeldungen | Abgeschlossen                                               | <a href="#">Ergebnisse anzeigen</a><br><a href="#">Ergebnisse herunterladen</a><br><a href="#">Daten exportieren</a>                                  |
| S1: Ausgewählte Probleme der pädagogischen Psychologie (Sommersemester 2019)   | 12 Einladungen<br>10 Rückmeldungen | Abgeschlossen                                               | <a href="#">Ergebnisse anzeigen</a><br><a href="#">Ergebnisse herunterladen</a><br><a href="#">Daten exportieren</a>                                  |

Impressum
AGB
Datenschutz

**Abbildung 1** Startseite von LEO für angemeldete Dozierende und verfügbare Funktionen.



Hohe Werte sprechen für eine sehr gute Bewertung: 1 (stimme gar nicht zu) bis 5 (stimme voll zu). Die Ergebnisse Ihrer Lehrveranstaltung sind als rote Linie dargestellt. Die graue Linie zeigt den Durchschnitt von 3660 Veranstaltungen in einer Vergleichsstichprobe. Für Vorlesungen, Seminare und Übungen zeigt die blaue Linie den Durchschnitt von nur Veranstaltung in diesem Format.

**Abbildung 2** Profillinien zum Vergleich der eigenen Lehrqualität mit anderen Veranstaltungen.