

Alternative evaluation methods in medical education

Marjo Wijnen-Meijer¹

1 TUD Dresden University of Technology, Medical Faculty and University Hospital Carl Gustav Carus, Institute of Medical Education, Dresden, Germany

Editorial

Quality assurance and the research and development of teaching are legally required and central components of higher education didactics. This is particularly true in medical education, which is shaped by specific demands and practice-oriented learning processes.

Feedback plays a crucial role in systematically capturing teaching quality, optimizing learning processes, and supporting both teachers and students in their development [1]. Two main types of feedback can be distinguished: feedback to teachers, which reflects their teaching quality and identifies areas for improvement, and feedback to students, which evaluates their performance and progress.

This article focuses on feedback to teachers, which is essential for the development of teaching and the maintenance of teaching quality.

Traditionally, universities use standardized student evaluations (SETs) in the form of standardized questionnaires, which are widely used to assess the quality of courses and teaching performance because of their simple handling and comparability [2]. However, these approaches often capture only snapshots and are limited in significance. In the medical context, where clinical exercises, simulations, and interprofessional learning play a central role, such methods often fail to adequately capture the complex teaching and learning processes. Numerous studies have shown that SETs are often strongly influenced by factors such as student expectations, gender, or ethnicity, and therefore may not always provide an objective assessment of teaching quality [3], [4]. At the same time, external factors such as course size or exam stress can influence the results [5], making them more of a “satisfaction analysis” rather than a true teaching evaluation [6].

In this context, the development and implementation of alternative evaluation approaches is becoming increasingly important. Methods such as realistic evaluation, as well as focus groups or peer feedback, allow for a more differentiated view of the contexts, dynamics, and effects in medical teaching situations.

The realistic evaluation, developed by Pawson and Tilley, focuses on understanding how programs work in specific contexts and the mechanisms that trigger these effects [7]. This method is particularly useful in medical education, as it not only examines outcomes but also the underlying conditions and processes. It is based on the principle: “What works, for whom, in what circumstances, and how?”. In teaching, this approach could represent a shift from purely summative assessments to analyzing the underlying learning processes and contexts. For example, one could investigate why a particular simulation training is successful for some students but not for others. The method requires both qualitative and quantitative data, such as interviews, observations, or performance assessments. Therefore, it is especially useful for studying complex teaching settings, as they occur in medical education.

In this context, realistic evaluation can be applied to analyze the effects of training programs, teaching methods or curriculum changes. The underlying core principles include mechanisms, contexts, and outcomes.

Mechanisms refer to the processes or underlying factors that explain how an intervention works. In a medical simulation, for instance, the mechanism might be the opportunity for learners to practice decision-making in a realistic environment.

Contexts involve the conditions or environments in which the intervention takes place. For example, learner engagement might depend on the availability of instructors or the cultural norms within the institution.

Finally, outcomes are the results or changes of the intervention. These might include improved clinical skills, enhanced teamwork, or increased self-confidence among participating students.

Insights from realistic evaluations help prepare instructional interventions to different types of learners and underlying contexts. In medical education, this could mean adapting teaching strategies for different groups, such as international students, or adjusting clinical instruction to the available resources at a hospital.

Peer observation can also be a valuable source of feedback, highlighting strengths and areas for development

in teaching [8]. This can be done using an observation guide, although it is not always necessary. The key is to clearly define the observational focus in advance.

Furthermore, focus groups can be a valuable addition to other evaluation methods in medical education, such as student evaluations. They provide an opportunity to collect qualitative student feedback and offer a more holistic perspective on teaching quality. However, the main disadvantage is that focus groups are time-consuming and require facilitation experience to ensure productive discussions [9]. Despite these challenges, focus groups can be a valuable tool for assessing and improving medical education, as they provide detailed, qualitative insights, encourage communication between students and instructors, and can contribute to more effective, student-centered teaching design [10].

Finally, student learning outcomes are also a useful feedback source for instructors [11]. This can be achieved through the analysis of exam results in the traditional sense, but also by using quiz questions during lectures to gain a current understanding of what students have comprehended and where knowledge gaps still persist. In this issue, various didactic interventions are evaluated. The article by Papan et al. describes the effect of an elective course on antimicrobial stewardship (AMS) [12]. The article by Ruck et al. evaluated the long-term effect of a smoking cessation counselling course using online questionnaires [13]. Tom Dreyer and colleagues compared the effect of two different teaching formats based on examination results [14]. Junga et al. compare the evaluations of compulsory and elective clerkships [15]. Brinkmann et al. have evaluated Mind-Body Medicine courses at two different universities using questionnaires and focus groups [16]. And Hofhansi et al. describe students' experiences with supervision, which were collected using questionnaires [17].

Enjoy reading this interesting issue!

Author's ORCID

Marjo Wijnen-Meijer: [0000-0001-8401-5047]

Competing interests

The author declares that she has no competing interests.

References

- Ding C, Sherman H. Teaching effectiveness and student achievement: Examining the relationship. *Educ Res Quart.* 2006;29(4):40-51.
- Hornstein HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Educ.* 2017;4(1):1304016. DOI: 10.1080/2331186X.2017.1304016
- Arnold IJM. Do examinations influence student evaluations? *Int J Educ Res.* 2009;48(4):215-224. DOI: 10.1016/j.ijer.2009.10.001
- Heffernan T. Sexism, racism, prejudice, and bias: A literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assess Eval High Educ.* 2022;47(1):144-154. DOI: 10.1080/02602938.2021.1888075,3
- O'Donovan R. Missing the forest for the trees: Investigating factors influencing student evaluations of teaching. *Assess Eval High Educ.* 2024;49(4):453-470. DOI: 10.1080/02602938.2023.2266862
- Rindermann H, Kohler J, Meisenberg G. Quality of Instruction Improved by Evaluation and Consultation of Instructors. *Int J Acad Dev.* 2007;12(2):73-85. DOI: 10.1080/13601440701604849
- Pawson R, Tilley N. *Realistic evaluation.* Thousand Oaks (CA): Sage; 1997.
- Georgiou H, Sharma M, Ling A. Peer review of teaching: What features matter? A case study within STEM faculties. *Innov Educ Teach Int.* 2018;55(2):190-200. DOI: 10.1080/14703297.2017.1342557
- Stalmeijer RE, McNaughton N, Van Mook WN. Using focus groups in medical education research: AMEE Guide No. 91. *Med Teach.* 2014;36(11):923-939. DOI: 10.3109/0142159X.2014.917165
- Hsieh KW, Iscoe MS, Lupton JR, Mains TE, Nayar SK, Orlando MS, Parzuchowski AS, Sabbagh MF, Schulz JC, Shenderov K, Simkin DJ, Vakili S, Vick JB, Xu T, Yin O, Goldberg HR. The Student Curriculum Review Team: How we catalyze curricular changes through a student-centered approach. *Med Teach.* 2015;37(11):1008-1012. DOI: 10.3109/0142159X.2014.990877
- Hattie J. Which Strategies Best Enhance Teaching and Learning in Higher Education? In: Mashek DJ, Hammer EY, editors. *Empirical Research in Teaching and Learning.* Hoboken (NJ): John Wiley & Sons, Ltd; 2011. p.130-142. DOI: 10.1002/9781444395341.ch8
- Papan C, Gärtner BC, Simon A, Müller R, Fischer MR, Darici D, Becker SL, Last K, Bushuven S. Stewards for Future: Piloting a medical undergraduate elective on antimicrobial stewardship. *GMS J Med Educ.* 2025;42(1):Doc9. DOI: 10.3205/zma001733
- Ruck J, Tiedemann E, Sudmann J, Kübler A, Simmenroth A. Evaluating the longitudinal effectiveness of a 5A model-based smoking cessation counselling course for medical students in family medicine placement. *GMS J Med Educ.* 2025;42(1):Doc10. DOI: 10.3205/zma001734
- Dreyer T, Papadopoulos S, Wiesner R, Karay Y. Classroom teaching versus online teaching in physiology practical course – does this lead to different examination results? *GMS J Med Educ.* 2025;42(1):Doc8. DOI: 10.3205/zma001732
- Junga A, Görlich D, Scherzer S, Schwarz M, Schulze H, Marschall B, Becker JC. “Hold the retractor, that's it?” – A retrospective longitudinal evaluation-study of the surgery and the elective tertial in the practical year. *GMS J Med Educ.* 2025;42(2):Doc3. DOI: 10.3205/zma001727
- Brinkmann B, Stöckigt B, Witt CM, Ortiz M, Herrmann M, Adam D, Vogelsang P. Reducing stress, strengthening resilience and self-care in medical students through Mind-Body Medicine (MBM). *GMS J Med Educ.* 2025;42(1):Doc7. DOI: 10.3205/zma001731
- Hofhansi A, Zlabinger G, Bach L, Röhrs J, Meyer AM, Rieder A, Wagner-Menghin M. Medical students' perception of supervision in MedUniVienna's structured internal medicine and surgery clerkship program: Subject specific differences and clerkship sequence effects. *GMS J Med Educ.* 2025;42(1):Doc5. DOI: 10.3205/zma001729

Corresponding author:

Prof. Dr. Marjo Wijnen-Meijer
TUD Dresden University of Technology, Medical Faculty
and University Hospital Carl Gustav Carus, Institute of
Medical Education, Fetscherstr. 74, D-01307 Dresden,
Germany
marjo.wijnen-meijer@tu-dresden.de

Please cite as

Wijnen-Meijer M. *Alternative evaluation methods in medical
education*. *GMS J Med Educ.* 2025;42(1):Doc15.
DOI: 10.3205/zma001739, URN: urn:nbn:de:0183-zma0017398

This article is freely available from
<https://doi.org/10.3205/zma001739>

Received: 2024-12-13
Revised: 2024-12-13
Accepted: 2024-12-13
Published: 2025-02-17

Copyright

©2025 Wijnen-Meijer. This is an Open Access article distributed under
the terms of the Creative Commons Attribution 4.0 License. See license
information at <http://creativecommons.org/licenses/by/4.0/>.

Alternative Evaluationsmethoden in der medizinischen Ausbildung

Marjo Wijnen-Meijer¹

¹ Technische Universität
Dresden, Medizinische
Fakultät und
Universitätsklinikum Carl
Gustav Carus, Institut für
Didaktik und Lehrforschung
in der Medizin, Dresden,
Deutschland

Leitartikel

Die Qualitätssicherung und Weiterentwicklung der Lehre ist ein rechtlich geforderter und zentraler Bestandteil der Hochschuldidaktik. Dies ist insbesondere in der durch komplexe Anforderungen und praxisorientierte Lernprozesse geprägten medizinischen Ausbildung der Fall. Feedback spielt hierbei eine entscheidende Rolle, um Lehrqualität systematisch zu erfassen, Lernprozesse zu optimieren und sowohl Lehrende als auch Lernende in ihrer Weiterentwicklung zu unterstützen [1]. Zwei zentrale Feedbackrichtungen lassen sich dabei unterscheiden: einerseits das Feedback an Lehrende, das deren Lehrqualität reflektiert und Verbesserungspotenziale aufzeigt und andererseits das Feedback an Studierende, welches deren Leistungen und Fortschritte bewertet.

Der Fokus dieses Artikels liegt auf dem Feedback an Lehrende, welches eine wesentliche Grundlage für die Weiterentwicklung der Lehre und die Sicherstellung der Lehrqualität bildet.

Traditionell verwenden Hochschulen standardisierte studentische Evaluationen (SETs), in Form von standardisierten Fragebögen, die durch ihre einfache Handhabung und Vergleichbarkeit eine breite Anwendung finden, um die Qualität der Lehrveranstaltungen und der Lehrleistung zu bewerten [2]. Diese Ansätze erfassen jedoch häufig nur Momentaufnahmen und bleiben in ihrer Aussagekraft begrenzt. Insbesondere im medizinischen Kontext, in dem klinische Übungen, Simulationen und interprofessionelles Lernen eine zentrale Rolle spielen, greifen solche Methoden oft zu kurz, um die komplexen Lehr- und Lernprozesse differenziert abbilden zu können. Zahlreiche Studien haben gezeigt, dass SETs oft stark von Faktoren, wie studentischen Erwartungen, Geschlecht oder ethnischer Zugehörigkeit beeinflusst werden und daher nicht immer eine objektive Einschätzung der Lehrqualität liefern können [3], [4]. Gleichzeitig können externe Faktoren wie

Kursgröße oder Prüfungsstress die Ergebnisse verzerren [5] und stellen somit eher eine „Zufriedenheitsanalyse“ als eine Lehrevaluation dar [6].

Vor diesem Hintergrund gewinnt die Entwicklung und Implementierung alternativer Evaluationsansätze zunehmend an Bedeutung. Methoden wie die realistische Evaluation, aber auch Fokusgruppen oder kollegiales Feedback ermöglichen eine differenzierte Betrachtung von den Kontexten, Dynamiken und Wirkungen in spezifischen Lehrsituationen.

Die realistische Evaluation, entwickelt von Pawson und Tilley, fokussiert sich darauf, wie Programme in spezifischen Kontexten wirken und welche Mechanismen diese Wirkungen auslösen [7]. Diese Methode ist besonders wertvoll in der medizinischen Lehre, da sie nicht nur die Ergebnisse, sondern auch die Rahmenbedingungen und zugrunde liegenden Prozesse untersucht. Sie beruht auf dem Grundprinzip: „What works, for whom, in what circumstances, and how?“. Die Anwendung in der Lehre könnte bedeuten, dass statt einer rein summativen Bewertung, die zugrundeliegenden Lernprozesse und Kontexte analysiert werden. Beispielsweise könnte man untersuchen, warum ein bestimmtes Simulationstraining bei bestimmten Studierenden erfolgreich ist und bei anderen nicht. Die Methodik erfordert qualitative und quantitative Daten, wie etwa Interviews, Beobachtungen oder Leistungsbewertungen. Deshalb eignet sie sich besonders für die Untersuchung komplexer Lehrsettings. Im Zusammenhang mit der medizinischen Ausbildung kann die sogenannte realistische Evaluation angewandt werden, um die Auswirkungen von Ausbildungsprogrammen, Lehrmethoden oder Lehrplanänderungen zu analysieren. Zu ihren Grundprinzipien gehören Mechanismen, Kontexte und Ergebnisse.

Mechanismen erfassen Prozesse oder zugrunde liegenden Faktoren, die erklären, wie eine Intervention funktioniert. Bei einer medizinischen Simulation könnte der Mechanismus zum Beispiel darin bestehen, dass die Lernenden die Möglichkeit haben, die Entscheidungsfindung in einer realistischen Umgebung zu üben.

Kontexte beinhalten hingegen die Bedingungen oder Umgebungen, in denen die Intervention wirkt und die ihren Erfolg beeinflussen können. So kann das Engagement der Lernenden beispielsweise von der Verfügbarkeit unterstützender Lehrkräfte oder den kulturellen Normen innerhalb der Einrichtung abhängen.

Schließlich werden auch die Ergebnisse oder Veränderungen, die durch die Intervention hervorgerufen werden, betrachtet. Dazu könnten verbesserte klinische Fertigkeiten, eine verstärkte Teamarbeit oder ein größeres Selbstvertrauen der Auszubildenden in der Medizin gehören.

Erkenntnisse aus realistischen Evaluierungen helfen dabei, didaktische Interventionen auf unterschiedliche Lernende und zugrundeliegenden Rahmenbedingungen zuzuschneiden. In der medizinischen Ausbildung könnte dies bedeuten, dass Lehrstrategien für unterschiedliche Gruppen, z. B. internationale Student*innen, angepasst werden oder dass der klinische Unterricht auf die Ressourcen des Krankenhauses zugeschnitten wird.

Die Fremdbeobachtung durch eine*n Kollegen*in (Peer) kann auch eine wertvolle Feedbackquelle für eigene Stärken und Entwicklungsmöglichkeiten in der Lehrtätigkeit sein [8]. Dies kann anhand eines Beobachtungsleitfadens erfolgen, ist jedoch nicht unbedingt notwendig. Wichtig ist, dass vorab der Fokus der Beobachtung geklärt wird.

Darüber hinaus können auch Fokusgruppen eine wertvolle Ergänzung zu anderen Bewertungsmethoden in der medizinischen Lehre sein, wie zum Beispiel den studentischen Evaluationen. Sie bieten die Möglichkeit, qualitatives Feedback von Studierenden zu sammeln und die Lehrqualität aus einer ganzheitlicheren Perspektive zu betrachten. Der Nachteil ist, dass diese zeitintensiv ist und Moderationserfahrung erfordert, um die Gespräche richtig zu führen [9]. Trotz dieser Herausforderungen können Fokusgruppen ein wertvolles Werkzeug zur Bewertung und Verbesserung der medizinischen Lehre sein, denn sie liefern detaillierte, qualitative Informationen, fördern die Kommunikation zwischen Studierenden und Dozierenden und können so zu einer effektiveren und studierendenzentrierten Lehrgestaltung beitragen [10]. Schließlich sind auch die Lernergebnisse von Student*innen eine gute Feedbackquelle für die Lehrenden [11]. Dies kann durch die Analyse von Prüfungsergebnissen im klassischen Sinne geschehen, aber auch durch Quizfragen in Vorlesungen, um eine aktuelle Vorstellung davon zu erhalten, was die Studierenden verstanden haben und wo noch Wissenslücken bestehen.

In diesem Heft werden verschiedene didaktische Interventionen evaluiert. Der Artikel von Papan et al. Beschreibt die Wirkung eines Wahlfachs zum Thema Antimicrobial Stewardship (AMS) [12]. Der Artikel von Ruck et

al. hat die Langzeitwirkung eines Raucherentwöhnungs-Beratungskurses mithilfe von Online-Fragebögen evaluiert [13]. Tom Dreyer und Kollegen haben die Wirkung von zwei verschiedenen Lehrformaten anhand von Prüfungsergebnissen verglichen [14]. Junga et al. vergleichen die Bewertungen von Pflicht- und Wahlterialen [15]. Brinkmann et al. haben Mind-Body Medicine Kurse an zwei verschiedenen Universitäten mithilfe von Fragebögen und Fokusgruppen evaluiert [16]. Und Hofhansi et al. beschreiben die Erfahrungen von Studierenden mit Supervision, die mithilfe von Fragebögen erfasst wurden [17].

Viel Spaß beim Lesen dieses interessanten Heftes!

ORCID der Autorin

Marjo Wijnen-Meijer: [0000-0001-8401-5047]

Interessenkonflikt

Die Autorin erklärt, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel hat.

Literatur

1. Ding C, Sherman H. Teaching effectiveness and student achievement: Examining the relationship. *Educ Res Quart.* 2006;29(4):40-51.
2. Hornstein HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Educ.* 2017;4(1):1304016. DOI: 10.1080/2331186X.2017.1304016
3. Arnold IJM. Do examinations influence student evaluations? *Int J Educ Res.* 2009;48(4):215-224. DOI: 10.1016/j.ijer.2009.10.001
4. Heffernan T. Sexism, racism, prejudice, and bias: A literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assess Eval High Educ.* 2022;47(1):144-154. DOI: 10.1080/02602938.2021.1888075,3
5. O'Donovan R. Missing the forest for the trees: Investigating factors influencing student evaluations of teaching. *Assess Eval High Educ.* 2024;49(4):453-470. DOI: 10.1080/02602938.2023.2266862
6. Rindermann H, Kohler J, Meisenberg G. Quality of Instruction Improved by Evaluation and Consultation of Instructors. *Int J Acad Dev.* 2007;12(2):73-85. DOI: 10.1080/13601440701604849
7. Pawson R, Tilley N. *Realistic evaluation.* Thousand Oaks (CA): Sage; 1997.
8. Georgiou H, Sharma M, Ling A. Peer review of teaching: What features matter? A case study within STEM faculties. *Innov Educ Teach Int.* 2018;55(2):190-200. DOI: 10.1080/14703297.2017.1342557
9. Stalmeijer RE, McNaughton N, Van Mook WN. Using focus groups in medical education research: AMEE Guide No. 91. *Med Teach.* 2014;36(11):923-939. DOI: 10.3109/0142159X.2014.917165

10. Hsieh KW, Iscoe MS, Lupton JR, Mains TE, Nayar SK, Orlando MS, Parzuchowski AS, Sabbagh MF, Schulz JC, Shenderov K, Simkin DJ, Vakili S, Vick JB, Xu T, Yin O, Goldberg HR. The Student Curriculum Review Team: How we catalyze curricular changes through a student-centered approach. *Med Teach*. 2015;37(11):10081012. DOI: 10.3109/0142159X.2014.990877
11. Hattie J. Which Strategies Best Enhance Teaching and Learning in Higher Education? In: Mashek DJ, Hammer EY, editors. *Empirical Research in Teaching and Learning*. Hoboken (NJ): John Wiley & Sons, Ltd; 2011. p.130-142. DOI: 10.1002/9781444395341.ch8
12. Papan C, Gärtner BC, Simon A, Müller R, Fischer MR, Darici D, Becker SL, Last K, Bushuven S. Stewards for Future: Piloting a medical undergraduate elective on antimicrobial stewardship. *GMS J Med Educ*. 2025;42(1):Doc9. DOI: 10.3205/zma001733
13. Ruck J, Tiedemann E, Sudmann J, Kübler A, Simmenroth A. Evaluating the longitudinal effectiveness of a 5A model-based smoking cessation counselling course for medical students in family medicine placement. *GMS J Med Educ*. 2025;42(1):Doc10. DOI: 10.3205/zma0011734
14. Dreyer T, Papadopoulos S, Wiesner R, Karay Y. Classroom teaching versus online teaching in physiology practical course – does this lead to different examination results? *GMS J Med Educ*. 2025;42(1):Doc8. DOI: 10.3205/zma001732
15. Junga A, Görlich D, Scherzer S, Schwarz M, Schulze H, Marschall B, Becker JC. “Hold the retractor, that's it?” – A retrospective longitudinal evaluation-study of the surgery and the elective tertiary in the practical year. *GMS J Med Educ*. 2025;42(2):Doc3. DOI: 10.3205/zma001727
16. Brinkmann B, Stöckigt B, Witt CM, Ortiz M, Herrmann M, Adam D, Vogelsang P. Reducing stress, strengthening resilience and self-care in medical students through Mind-Body Medicine (MBM). *GMS J Med Educ*. 2025;42(1):Doc7. DOI: 10.3205/zma001731
17. Hofhansl A, Zlabinger G, Bach L, Röhrs J, Meyer AM, Rieder A, Wagner-Menghin M. Medical students' perception of supervision in MedUniVienna's structured internal medicine and surgery clerkship program: Subject specific differences and clerkship sequence effects. *GMS J Med Educ*. 2025;42(1):Doc5. DOI: 10.3205/zma001729

Korrespondenzadresse:

Prof. Dr. Marjo Wijnen-Meijer
Technische Universität Dresden, Medizinische Fakultät
und Universitätsklinikum Carl Gustav Carus, Institut für
Didaktik und Lehrforschung in der Medizin, Fetscherstr.
74, 01307 Dresden, Deutschland
marjo.wijnen-meijer@tu-dresden.de

Bitte zitieren als

Wijnen-Meijer M. *Alternative evaluation methods in medical education*. *GMS J Med Educ*. 2025;42(1):Doc15.
DOI: 10.3205/zma001739, URN: urn:nbn:de:0183-zma0017398

Artikel online frei zugänglich unter

<https://doi.org/10.3205/zma001739>

Eingereicht: 13.12.2024

Überarbeitet: 13.12.2024

Angenommen: 13.12.2024

Veröffentlicht: 17.02.2025

Copyright

©2025 Wijnen-Meijer. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.