

Article

Educator Feedback Skill Assessment: An Educational Survey Design Study

Alex Moroz ^{1,*}, Jennifer Stone ², Francis Lopez ¹, Cynthia Racine ³ and Kristin Carmody ⁴¹ Department of Rehabilitation Medicine, NYU Grossman School of Medicine, New York, NY 10016, USA² New York Medical College, Valhalla, NY 10595, USA³ Rehabilitation Medicine Service, Robley Rex VA Hospital, Louisville, KY 40206, USA⁴ Kristin Carmody Emergency Medicine, New York, NY 10016, USA

* Correspondence: alex.moroz@nyulangone.org; Tel.: +1-212-263-6110

Abstract: Background: Delivering impactful feedback is a skill that is difficult to measure. To date there is no generalizable assessment instrument which measures the quality of medical education feedback. The purpose of the present study was to create an instrument for measuring educator feedback skills. Methods: Building on pilot work, we refined an assessment instrument and addressed content and construct validity using expert validation (qualitative and quantitative). This was followed by cognitive interviews of faculty from several clinical departments, which were transcribed and analyzed using ATLAS.ti qualitative software. A research team revised and improved the assessment instrument. Results: Expert validation and cognitive interviews resulted in the Educator Feedback Skills Assessment, a scale with 10 items and three response options for each. Conclusions: Building on the contemporary medical education literature and empiric pilot work, we created and refined an assessment instrument for measuring educator feedback skills. We also started the argument on validity and addressed content validity.

Keywords: faculty development; educational assessment; graduate medical education



Citation: Moroz, A.; Stone, J.; Lopez, F.; Racine, C.; Carmody, K. Educator Feedback Skill Assessment: An Educational Survey Design Study. *Int. Med. Educ.* **2022**, *1*, 97–105. <https://doi.org/10.3390/ime1020012>

Academic Editor: Hideki Kasuya

Received: 28 November 2022

Accepted: 2 December 2022

Published: 9 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Conceptual Framework

The ultimate goal of assessment practices in professional health education is improved healthcare. High-quality and credible feedback is necessary to provide a meaningful mechanism through which physicians can be expected to grow [1]. Feedback is fundamental for everything we do—it is an essential part of every framework, every curriculum, and every teaching interaction.

Despite the importance of feedback, residents and faculty thought that provider feedback skills were not sufficiently developed [2,3]. Similarly, faculty from both university and community-based programs described having minimal training and a lack of understanding of the best practices for delivering feedback [4], despite the availability of excellent practical guides [5–7]. It does not appear that this is a perception issue—a qualitative study of simulated feedback encounters suggested that faculty skills do not match recommended practice in a number of areas [8].

There is growing evidence that utilizing teacher-centered models of feedback is not sufficient to improving the quality of feedback [9–14]. Characteristics of feedback providers form one of the three clusters seen when viewing feedback through the lens of the socio-cultural model [15]. For example, improving feedback provider skills may in fact improve outcomes. Sargeant and other colleagues have shown that training coaches to conduct a reflective feedback conversation can improve the acceptance and uptake of feedback [16]. Similarly, supportive coaching has been associated with both perceived coach competence and satisfaction in the sports realm [17].

1.2. Related Research

In order to explore the intended meaning and breadth of the feedback construct, we completed the following steps in a pilot study [18]. We started by conducting a literature review that aligned the feedback construct with prior research and identified existing feedback scales. We then explored how feedback participants conceptualize and describe feedback. We asked feedback recipients (resident physicians) to select, script, and enact six faculty-resident feedback vignettes.

We then conducted seven faculty focus groups that included 23 feedback providers. We asked the faculty, who watched each vignette video as a group, to comment on elements that were successful and on areas for improvement. Synthesizing the literature review and focus group findings ensured that our conceptualization of the feedback construct made theoretical sense to scholars in the field and used language that feedback providers understood. It allowed us to draft a list of 51 items that we grouped under 10 proposed dimensions of feedback and to create an early assessment scale, initially named Feedback Rating Scale (Appendix A Table A1).

Although several feedback delivery frameworks have been described, these are applicable to narrow areas within medical education. Several assessments were developed within specific contexts—written feedback [19], simulation debriefing [20], direct observation of clinical skills [21], communication skills feedback [22], feedback by residents [23], and feedback assessed by medical students [24,25]—however, these instruments are not generalizable to other types of feedback. The major research gap in this domain is therefore the absence of a reliable measurement instrument that can be applied to multiple facets of medical education.

The purpose of the present study was to (a) define dimensions that best represent the construct of feedback in medical education, and to (b) create and refine a generalizable assessment instrument for measuring educator feedback skills.

2. Materials and Methods

2.1. Research Model

This is an educational survey design study. We adopted Messick's construct validity framework [26]. We selected Messick's framework because, in contrast to the earlier validity frameworks that focused on "types" of validity (e.g., content or criterion), this approach favors a unified framework in which construct validity (the only type) is supported by evidence derived from multiple sources [27]. We envisioned our study findings being one of such sources that begin the "validity argument".

For additional guidance in the study design, we selected a systematic and practical approach for creating high-quality survey scales that synthesized multiple survey design techniques into a cohesive mixed-methods process [28]. Building on our pilot work, we addressed the content, construct, and response process aspects of validity.

2.2. Participants

1. To explore the content aspect of construct validity using expert validation, we recruited an international panel of methodologists, researchers, and subject-matter experts.
2. To conduct cognitive interviews, we recruited experienced feedback providers from 4 clinical departments (Emergency Medicine, Medicine, Orthopedic Surgery, Physical Medicine and Rehabilitation) at a single academic health system.

2.3. Data Collection Tools

1. The experts were asked to comment on each item's representativeness, clarity, relevance and distribution using an anonymous online form: https://docs.google.com/forms/d/e/1FAIpQLSffLngxbC_XTBv31dQDi0ftczjz3wDMGrfz_ZcOmLimcnPXiA/viewform (accessed on 5 December 2022).
2. Experts rated each item as essential, useful but not essential, or not necessary using an anonymous online form: <https://docs.google.com/forms/d/e/1FAIpQLSfd8>

3pEZhq_z-KRhDLFNM3bJCtXfRopCFAIRAb_TTs1D96J0g/viewform?usp=sf_link (accessed on 5 December 2022).

2.4. Data Collection Process

To assess how clear and relevant the items are with respect to the construct of interest, international experts were asked to comment on each item's representativeness, clarity, relevance, and distribution using an anonymous online form. We also asked the experts to review the labels used for the response categories (qualitative review: content aspect of construct validity using expert validation). We asked the same group of experts to review individual items in the modified assessment instrument. Experts rated each item as essential, useful but not essential, or not necessary using an anonymous online form (quantitative review: content aspect of construct validity using expert validation).

To ensure that the respondents interpreted items as we intended (response process validity) we asked experienced feedback providers to use the assessment instrument, modified in above steps, and to rate videotaped feedback encounters that we developed as part of the pilot study [18]. We then conducted structured individual cognitive interviews utilizing the technique of concurrent probing [29]. This technique involves the interviewer asking about the respondent's thought process while they are completing the questionnaire, and allows a reasonable balance between the demand on the respondent and minimizing the recall bias [28].

2.5. Data Analysis

Using the data collected during the qualitative reviews, we used expert responses and comments to modify and revise the assessment instrument. During quantitative expert reviews, we used both a predetermined content validity ratio cut-point (McKenzie recommends 0.62 minimum for statistical significance of <0.05 for a 10-member panel), and the narrative comments by experts to make inclusion and exclusion decisions for individual items [30].

Audio files of recorded cognitive interviews were transcribed, coded, and analyzed qualitatively using the ATLAS.ti software (Scientific Software Development GmbH 2019) in order to modify and improve the overall assessment instrument and the individual survey items. The research team used a consensus method in deciding on whether to proceed with each revision suggested by interviewees; suggestions that received at least three out of four research team votes were implemented.

3. Results

The majority of interviews were conducted face to face, however, the last two interviews were done virtually due to the COVID-19 pandemic-related restrictions. The assessment instrument (final version, Appendix A Table A2) was revised eight times during the research study (Table 1). The instrument name was changed from Feedback Rating Scale to Educator Feedback Skill Assessment (EFSA).

Qualitative review. Twelve experts agreed to participate (see Acknowledgements section). Ten of the twelve submitted narrative comments online. In addition to individual item revisions, the number of items was increased from 31 to 32 (one item was split to avoid "double barreling").

Quantitative review. Eight of the twelve experts submitted "inclusion/exclusion" votes online. Ten of the thirty-two items had the content validity ratios of > 0.62 , and were included in the final version of the assessment instrument.

Cognitive interviews. Twelve cognitive interviews were conducted, ten face to face and two online via Zoom. Participants included four teaching faculty in Emergency Medicine, four in Physical Medicine and Rehabilitation, three in Internal Medicine, and one in Orthopedic Surgery. Qualitative analysis of the interview transcripts yielded twenty-three recommendations. Seven of the suggestions received at least three out of four research team votes and were implemented in the final version of EFSA. To arrive at the final version

of the assessment instrument (Appendix A Table A2), the PI made several additional changes to improve readability, reduce wordiness, and improve item format consistency.

Table 1. Tabulated Results.

| | | |
|-----------------------------|---|----------------------------|
| Expert Qualitative Reviews | Experts Recruited | 12 |
| | Expert Comments | 10 |
| | Item Number Changes | Increased by 1 (31 to 32) |
| Expert Quantitative Reviews | Experts Recruited | 12 |
| | Expert Votes | 8 |
| | Item Number Changes | Decreased by 22 (32 to 10) |
| Cognitive Interviews | Total Participants Recruited | 12 |
| | Participants in Emergency Medicine | 4 |
| | Participants in Physical Medicine | 4 |
| | Participants in Internal Medicine | 3 |
| | Participants in Orthopedic Surgery | 1 |
| | Total Recommendations | 23 |
| | Incorporated Recommendations (>75% Votes) | 7 |
| Instrument Revisions | | 8 |
| Instrument Name Changes | | 1 |

4. Discussion

We believe a rigorous instrument that builds on existing theory and empirical evidence is necessary to measure the quality of feedback in medical education. Our study takes the first step in creating and validating such an instrument. Our results may also impact assessment in medical education in several ways.

Firstly, our findings may deepen the theoretical understanding of the dimensions of feedback necessary for making it meaningful and impactful, with the potential benefit for both medical education researchers and practitioners. Secondly, defining performance expectations for feedback providers in the form of a practical rubric can enhance reliable scoring of feedback performance assessments. Finally, although rubrics may not facilitate valid judgment of feedback assessments per se, they have the potential of promoting learning and improving the instruction of feedback providers by making expectations and criteria explicit, thereby facilitating feedback and self-assessment [31].

Our work started from de novo observations of feedback in a pilot project. While our findings were undoubtedly colored by the work of others and existing frameworks for feedback, we expect to further validate current methods of assessment, and explore and define novel dimensions of delivering feedback. Our work also built on an emerging area of feedback research supported by recent work of others and by our pilot work: specificity. Roze des Ordon and other colleagues identified variability in feedback recipients (4 ‘resident challenges’) and suggested adjusting feedback provider approaches accordingly [8]. Our own pilot [18], on the other hand, was based on scenarios that were selected, scripted, and enacted by learners (resident physicians), and the resultant data suggested additional variability in feedback providers. Using more than one perspective in developing items and dimensions of the assessment instrument may allow us to highlight multiple facets of the feedback construct and understand it more fully.

We think that the collaborative nature of this study is also a strength. Several prominent scholars with unique knowledge in assessment and feedback agreed to participate in expert validation (see Acknowledgements section). Within our own institution, we included faculty from 4 diverse departments in the cognitive interviewing, from both “cognitive” and “procedural” specialties, which supports the generalizability of the resultant instrument.

We addressed only one (content) of the four aspects (structural, content, generalizability, and substantive) of validity described by Messick [26], and this is undoubtedly the greatest weakness of this work. However, we feel strongly that once our new instrument is available to the medical education research community, the sooner this shortcoming can be addressed, by ourselves and by others. Additionally, early use of the instrument by the medical educators in the field is likely to provide feedback that will allow us to further refine and polish EFSA. Future studies will need to explore multiple facets of the feedback construct, while varying the types of feedback providers and feedback recipients. Another area of interest involves the study of different relationship stages, for example, one-time feedback vs ongoing coaching, 'on the fly' vs scheduled at the end of a clinical rotation.

To continue collecting the validity evidence, future studies should delve into the psychometric properties of EFSA focusing on structural aspects, as well as convergent and discriminant validity (external aspect). Future studies should also explore the relationship between EFSA and additional external measures such as motivation to use feedback, feedback-seeking frequency, and satisfaction with feedback using existing survey items [32]. The change in physicians' behavior and performance and how they affect patient outcomes are also areas of future interest. Additional studies across different specialty areas and demographic variables should also be conducted to further explore the generalizability aspect of construct validity.

5. Conclusions

Building on the contemporary medical education literature and empiric pilot work, we created and refined an assessment instrument for measuring educator feedback skills. We also started the argument on validity and addressed content validity. Future studies should address structural, generalizability, and substantive aspects of validity, and test the new instrument in a variety of settings and contexts.

Author Contributions: A.M. contributed to study planning, data collection and analysis, manuscript writing. J.S. contributed to data collection, manuscript writing. F.L. contributed to study planning, data analysis, manuscript writing. C.R. contributed to study planning, data analysis, manuscript writing. K.C. contributed to study planning, data analysis, manuscript writing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Berenstein Foundation and Program for Medical Education Innovations and Research (PrMEIR) at NYU Grossman School of Medicine.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of NYU Grossman School of Medicine.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank our jurors for their time and dedication—Anthony Artino, Eric Warm, Heather Armson, Lisa Altschuler, Eric Holmboe, Pim Teunissen, Amanda Lee Roze des Ordons, Sondra Zabar, Adina Kalet, Jeremy Branzetti, Donna Phillips, David Stern.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Pilot Feedback Rating Scale (FP = Feedback-Provider, FR = Feedback-Recipient).

| Feedback Dimension | Feedback Items | | | | | | |
|-------------------------------------|---|--|-------------------|-------------------|----------------|----------------|----------------|
| Preparation, Engagement, Investment | FP dedicated adequate time to the feedback conversation | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP was honest about not enough time or not enough facts | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP ensured quiet, private, appropriate environment | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP minimized disruptions | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP was prepared, present, engaged and paying attention | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| Defining Expectations | FP was making eye contact and leaning forward | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP was not 'just going through the motions' | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP was organized and completed the encounter | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP defined expectations for performance | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | Encouraging Self-Assessment | FP encouraged the FR to self-assess | | | | | |
| | | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| Beneficence, Encouragement, Respect | FP was warm, approachable, supportive, encouraging, & reassuring | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP was positive and used positive language | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP was polite and respectful | | | | | | |
| Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | | |
| Exploration, Reaction, Dialogue | FP was constructive without offending | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP listened | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP facilitated a dialogue | | | | | | |
| Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | | |
| | FP reacted to FR self-assessment and other comments | | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |
| | FP probed deeper and asked for elaboration | | | | | | |
| Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | | |

Table A1. Cont.

| Feedback Dimension | Feedback Items | | | | | |
|-----------------------------------|--|-------------------|-------------------|----------------|----------------|----------------|
| Using Facts and Observations | Feedback was based on observed performance by FR | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| Specificity, Use of Examples | FP described specific examples of specific FR behaviors | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| Confidence, Direction, Correction | FP remained calm, composed, and non-confrontational | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| | FP redirected and disarmed | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| | FP confronted wrong resident perceptions | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| Individualizing Conversation | FP confronted inappropriate FR behaviors | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| | FP adapted the feedback conversation and their approach based on FR comments and behaviors during the feedback encounter | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| | Feedback conversation included specific areas for improvement | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| Next Steps | Feedback conversation included measurable goals | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| | Feedback conversation included realistic action plan | | | | | |
| | Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree |
| | Feedback conversation included discussion of a timely follow-up | | | | | |
| Strongly Disagree | Disagree | Somewhat Disagree | Somewhat Agree | Agree | Strongly Agree | |

Table A2. Educator Feedback Skills Assessment.

| Items | Rating | | | Comments |
|---|--|---|---|----------|
| Educator appeared engaged | Distracted | Inconsistently engaged | Consistently engaged | |
| Educator was prepared for the feedback session | Unprepared for the feedback session; did not know Learner or his/her performance | Prepared for the feedback session; knew some things about Learner and his/her performance | Prepared for the feedback session; knew Learner and his/her performance in detail | |
| Self-assessment encouraged and incorporated in conversation | Self-assessment neither encouraged nor incorporated in conversation | Self-assessment encouraged OR incorporated in conversation | Self-assessment encouraged AND incorporated in conversation | |
| Educator was respectful | Disrespectful | Inconsistently respectful | Consistently respectful | |
| Educator was constructive | Not constructive | Inconsistently constructive | Consistently constructive | |

Table A2. Cont.

| Items | | Rating | Comments |
|---|---|---|---|
| Educator facilitated dialogue | Did not ask questions, did not allow time for or dismissed Learner comments | Asked some questions, reacted to Learner comments | Asked many questions, allowed time for responses, encouraged Learner comments |
| Educator probed deeper and asked for elaboration | Did not ask for clarification or elaboration | Inconsistently asked for clarification or elaboration | Consistently asked for clarification or elaboration |
| Educator provided specific examples to Learner | Educator provided no examples | Educator provided at least one specific example | Educator provided many specific examples |
| Conversation included specific areas for improvement (WHAT to improve) | Conversation did not include areas for improvement | Conversation included at least one area for improvement | Conversation included many areas for improvement |
| Conversation included an action plan (HOW to improve) | Action plan was not discussed | Action plan was discussed in general terms | A specific action plan was discussed |
| GENERAL COMMENTS/ADVICE Please include any suggestions for this Educator | | | |

References

- Eva, K.W.; Armson, H.; Holmboe, E.; Lockyer, J.; Loney, E.; Mann, K.; Sargeant, J. Factors influencing responsiveness to feedback: On the interplay between fear, confidence, and reasoning processes. *Adv. Health Sci. Educ.* **2011**, *17*, 15–26. [[CrossRef](#)]
- Carmody, K.; Walia, I.; Coneybeare, D.; Kalet, A. Can a Leopard Change Its Spots? A Mixed Methods Study Exploring Emergency Medicine Faculty Perceptions of Feedback, Strategies for Coping and Barriers to Change. Master's Thesis, Maastricht University School of Health Education, Maastricht, The Netherlands, 2017.
- Moroz, A.; Horlick, M.; Mandalaywala, N.; Stern, D.T. Faculty feedback that begins with resident self-assessment: Motivation is the key to success. *Med. Educ.* **2018**, *52*, 314–323. [[CrossRef](#)]
- Kogan, J.R.; Conforti, L.N.; Bernabeo, E.C.; Durning, S.J.; Hauer, K.E.; Holmboe, E.S. Faculty staff perceptions of feedback to residents after direct observation of clinical skills. *Med. Educ.* **2012**, *46*, 201–215. [[CrossRef](#)]
- Lefroy, J.; Watling, C.; Teunissen, P.; Brand, P.L.P. Guidelines: The do's, don'ts and don't knows of feedback for clinical education. *Perspect. Med. Educ.* **2015**, *4*, 284–299. [[CrossRef](#)]
- Roze des Ordon, A.L.; Gaudet, J.; Grant, V.; Harrison, A.; Millar, K.; Lord, J. Clinical feedback and coaching—BE-SMART. *Clin. Teach.* **2019**, *17*, 255–260. [[CrossRef](#)]
- Sargeant, J.; Lockyer, J.M.; Mann, K.; Armson, H.; Warren, A.; Zetkolic, M.; Soklaridis, S.; Könings, K.D.; Ross, K.; Silver, I.; et al. The R2C2 model in residency education: How does it foster coaching and promote feedback use? *Acad. Med.* **2018**, *93*, 1055–1063. [[CrossRef](#)]
- Roze des Ordon, A.L.; Cheng, A.; Gaudet, J.E.; Downar, J.; Lockyer, J.M. Exploring Faculty Approaches to Feedback in the Simulated Setting. *Simul. Health J. Soc. Simul. Healthc.* **2018**, *13*, 195–200. [[CrossRef](#)]
- Bing-You, R.; Hayes, V.; Varaklis, K.; Trowbridge, R.; Kemp, H.; McKelvy, D. Feedback for Learners in Medical Education: What is Known? A Scoping Review. *Acad. Med.* **2017**, *92*, 1346–1354. [[CrossRef](#)]
- Bing-You, R.; Ramani, S.; Ramesh, S.; Hayes, V.; Varaklis, K.; Ward, D.; Blanco, M. The interplay between residency program culture and feedback culture: A cross-sectional study exploring perceptions of residents at three institutions. *Med. Educ. Online* **2019**, *24*, 1611296. [[CrossRef](#)]
- Bing-You, R.G.; Trowbridge, R.L. Why Medical Educators May Be Failing at Feedback. *JAMA* **2009**, *302*, 1330–1331. [[CrossRef](#)]
- Kraut, A.; Yarris, L.M.; Sargeant, J. Feedback: Cultivating a Positive Culture. *J. Grad. Med. Educ.* **2015**, *7*, 262–264. [[CrossRef](#)]
- Molloy, E.; Ajjawi, R.; Bearman, M.; Noble, C.; Rudland, J.; Ryan, A. Challenging feedback myths: Values, learner involvement and promoting effects beyond the immediate task. *Med. Educ.* **2019**, *54*, 33–39. [[CrossRef](#)]
- Telio, S.; Ajjawi, R.; Regehr, G. The "Educational Alliance" as a Framework for Reconceptualizing Feedback in Medical Education. *Acad. Med.* **2015**, *90*, 609–614. [[CrossRef](#)]
- Ramani, S.; Könings, K.D.; Ginsburg, S.; van der Vleuten, C.P. Feedback Redefined: Principles and Practice. *J. Gen. Intern. Med.* **2019**, *34*, 744–749. [[CrossRef](#)]

16. Sargeant, J.; Lockyer, J.; Mann, K.; Holmboe, E.; Silver, I.; Armson, H.; Driessen, E.; MacLeod, T.; Yen, W.; Ross, K.; et al. Facilitated reflective performance feedback: Developing an evidence- and theory-based model that builds relationship, explores reactions and content, and coaches for performance change (R2C2). *Acad. Med.* **2015**, *90*, 1698–1706. [[CrossRef](#)]
17. Pulido, J.J.; García-Calvo, T.; Leo, F.M.; Figueiredo, A.J.; Sarmiento, H.; Sánchez-Oliva, D. Perceived coach interpersonal style and basic psychological needs as antecedents of athlete-perceived coaching competency and satisfaction with the coach: A multi-level analysis. *Sport Exerc. Perform. Psychol.* **2020**, *9*, 16–28. [[CrossRef](#)]
18. Moroz, A.; King, A.; Kim, B.; Fusco, H.; Carmody, K. Constructing a Shared Mental Model for Feedback Conversations: Faculty Workshop Using Video Vignettes Developed by Residents. *MedEdPORTAL* **2019**, *15*, 10821. [[CrossRef](#)]
19. Warm, E.; Kelleher, M.; Kinnear, B.; Sall, D. Feedback on Feedback as a Faculty Development Tool. *J. Grad. Med. Educ.* **2018**, *10*, 354–355. [[CrossRef](#)]
20. Minehart, R.D.; Rudolph, J.; Pian-Smith, M.C.M.; Raemer, D.B. Improving Faculty Feedback to Resident Trainees during a Simulated Case. *Anesthesiology* **2014**, *120*, 160–171. [[CrossRef](#)]
21. Halman, S.; Dudek, N.; Wood, T.; Pugh, D.; Touchie, C.; McAleer, S.; Humphrey-Murto, S. Direct Observation of Clinical Skills Feedback Scale: Development and Validity Evidence. *Teach. Learn. Med.* **2016**, *28*, 385–394. [[CrossRef](#)]
22. Perron, N.J.; Nendaz, M.; Louis-Simonet, M.; Sommer, J.; Gut, A.; Baroffio, A.; Dolmans, D.; van der Vleuten, C. Effectiveness of a training program in supervisors' ability to provide feedback on residents' communication skills. *Adv. Health Sci. Educ.* **2012**, *18*, 901–915. [[CrossRef](#)] [[PubMed](#)]
23. Bashir, K.; Elmoheen, A.; Seif, M.; Anjum, S.; Farook, S.; Thomas, S. In Pursuit of the Most Effective Method of Teaching Feedback Skills to Emergency Medicine Residents in Qatar: A Mixed Design. *Cureus* **2020**, *12*, e8155. [[CrossRef](#)] [[PubMed](#)]
24. Bing-You, R.; Ramesh, S.; Hayes, V.; Varaklis, K.; Ward, D.; Blanco, M. Trainees' Perceptions of Feedback: Validity Evidence for Two FEEDME (Feedback in Medical Education) Instruments. *Teach. Learn. Med.* **2018**, *30*, 162–172. [[CrossRef](#)]
25. Richard-Lepouriel, H.; Bajwa, N.; De Grasset, J.; Audétat, M.; Dao, M.D.; Jastrow, N.; Nendaz, M.; Perron, N.J. Medical students as feedback assessors in a faculty development program: Implications for the future. *Med. Teach.* **2020**, *42*, 536–542. [[CrossRef](#)] [[PubMed](#)]
26. Messick, S. Validity of Psychological Assessment. *Am. Psychol.* **1995**, *50*, 741–749. [[CrossRef](#)]
27. A Cook, D.A.; Brydges, R.; Ginsburg, S.; Hatala, R. A contemporary approach to validity arguments: A practical guide to Kane's framework. *Med. Educ.* **2015**, *49*, 560–575. [[CrossRef](#)] [[PubMed](#)]
28. Artino, A.R., Jr.; La Rochelle, J.S.; DeZee, K.J.; Gehlbach, H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med. Teach.* **2014**, *36*, 463–474. [[CrossRef](#)]
29. Watt, T.; Rasmussen, Å.K.; Groenvold, M.; Bjorner, J.B.; Watt, S.H.; Bonnema, S.J.; Hegedüs, L.; Feldt-Rasmussen, U. Improving a newly developed patient-reported outcome for thyroid patients, using cognitive interviewing. *Qual. Life Res.* **2008**, *17*, 1009–1017. [[CrossRef](#)]
30. McKenzie, J.; Wood, M.; Kotecki, J.; Clark, J.; Brey, R. Establishing content validity: Using qualitative and quantitative steps. *Am. J. Health Behav.* **1999**, *23*, 311–318. [[CrossRef](#)]
31. Jonsson, A.; Svingby, G. The use of scoring rubrics: Reliability, validity and educational consequences. *Educ. Res. Rev.* **2007**, *2*, 130–144. [[CrossRef](#)]
32. Steelman, L.; Levy, P.E.; Snell, A.F. The Feedback Environment Scale: Construct Definition, Measurement, and Validation. *Educ. Psychol. Meas.* **2004**, *64*, 165–184. [[CrossRef](#)]