ORIGINAL RESEARCH

Feedback Quality in Geriatric Medicine: Analyzing Entrustable Professional Activities in a Competency-Based Curriculum

Azin Dolatabadi, BA&Sc[1], Aathmika Nandan, HBSc[1], Dov Gandell, MDCM[2]

[1]*Temerty Faculty of Medicine, University of Toronto, Toronto;* [2]*Sunnybrook Health Sciences Centre, Department of Geriatric Medicine, Toronto, ON*

https://doi.org/10.5770/cgj.28.848

ABSTRACT

Background

Competency-based medical education (CBME) aims to enhance the quality of medical training by providing timely, actionable feedback through entrustable professional activities (EPAs). However, variability in feedback quality remains a concern across residency programs.

Methods

We conducted a retrospective analysis of EPA feedback forms from a geriatric medicine program, comparing two distinct time periods: 2019–2020 and 2021–2022. This program averages eight residents per year with 30 full-time faculty members. The quality of feedback was assessed based on timeliness, task orientation, actionability, and polarity.

Results

404 EPAs were initiated and completed in 2019–2020, with 69% (n=278) being timely, 89% (n=360) task oriented, 50% (n=203) very actionable, and 62% (n=250) corrective in polarity. 851 EPAs were initiated in 2021–2022 and 76% (n=645) were completed, with 64% (n=410) being timely, 78% (n=501) task oriented, 40% (n=259) very actionable, and 47% (n=305) corrective in polarity. Timely feedback was more likely to be task-oriented ($\chi^2(1)=11.87$, $p<.001$), actionable ($\chi^2(2)=24.40$, $p<.001$), and corrective ($\chi^2(3)=22.80$, $p<.001$) in the second timeframe. Compared to the second timeframe, EPAs completed in the first timeframe were more likely to be task oriented ($\chi^2(1)=22.08$, $p<.001$), actionable ($\chi^2(2)=25.54$, $p<.001$), and corrective in polarity ($\chi^2(3)=25.89$, $p<.001$).

Conclusions

Our study revealed lower quality feedback over time since implementing CBME at a geriatric medicine subspecialty training program. The root causes of the reduction in quality were not explored but are theorized to be multifactorial. Further investigation into the reasons for the reduction in feedback quality will help direct interventions to better sustain the quality of CBME implementation.

Key words: feedback quality, entrustable professional activities (EPAs), geriatric medicine, competency-based medical education (CBME), postgraduate medical education

INTRODUCTION

In recent years, residency training in Canada has significantly evolved with the adoption of competency-based medical education (CBME). This shift aimed to better align educational programs with the competencies required for effective clinical practice, thereby enhancing clinician preparedness to meet patient needs.[1] Unlike traditional time-based curricula, CBME emphasizes ongoing competency assessment, accountability, and learner-centered outcomes.[1,2] The core of CBME is the use of entrustable professional activities (EPAs), which are specialty-specific clinical tasks that residents must perform competently and independently to progress in their training.[3,4] The term EPA describes both the specific clinical tasks and the assessments that contextualize them.

EPAs are designed to provide structured, formative, and timely feedback, focusing on detailed observations of clinical practice for learner development.[5] Unlike summative traditional in-training evaluation reports (ITERs) at the end of rotations, EPAs emphasize real-time, task-specific assessments, promoting continuous learning and improvement.[6,7] Despite these advantages, the effectiveness of EPAs in delivering high-quality feedback remains uncertain, prompting further investigation into their implementation and impact.

The quality of feedback is crucial for medical education as it supports the acquisition of new knowledge and skills, which is essential given the shift towards CBME.[8] However, studies have highlighted challenges in achieving

effective feedback in clinical settings, including the lack of timely, specific, and actionable feedback.[9,10] For instance, studies of end-rotation evaluations have shown that feedback tends to be vague and not directly linked to observed clinical behaviors, which undermines its utility for learning.[7,11] Conversely, EPAs, which are based on direct observation of clinical encounters, are expected to provide more immediate and detailed feedback.[12]

Despite the theoretical advantages of CBME and EPAs, the practical implementation of these concepts has revealed several challenges.[13,14] For example, research from an internal medicine program found feedback to be task-oriented, but often lacking timeliness and actionable recommendations.[15] Similar findings were reported in a smaller oncology training program where only about one-third of EPA feedback was actionable.[5] Differences may also exist between residency programs, as a recent study found that smaller programs and those with an earlier CBME launch year offer higher quality feedback to learners.[16]

Given these challenges, ongoing evaluation of feedback quality within EPAs is critical to identify trends and areas for improvement in EPA feedback as part of continuous quality improvement. The purpose of this study was to assess the quality of written feedback documented in EPAs for postgraduate year (PGY)-4 and -5 geriatric subspecialty residents at the University of Toronto. Building on the framework established by Madrazo *et al*,[15] we examined feedback in terms of timeliness, task orientation, polarity, and actionability across two time points: 2019–2020, the first year of EPA implementation, and 2021–2022, two years later.

METHODS

Setting

The University of Toronto's Geriatric Medicine subspecialty training program in Toronto, Ontario, Canada, implemented Competence By Design (CBD; The Royal College of Physicians and Surgeons' of Canada's version of CBME) on July 1st, 2019. At this institution, EPAs are completed electronically by assessors using Elentra, an online platform (https//elentra.com) accessible via computers or mobile devices. Starting in 2021–2022, EPAs were given an expiry limit of two weeks from the date of initiation, with an opportunity to reset a single time for an additional week. Resets can only be initiated by the program administrator, the program director, or an Elentra Systems administrator. Residents are required to be assessed on a specific number of each EPA as stipulated by the Royal College of Physicians and Surgeons of Canada. Assessments can only be completed by faculty members, clinical associates, or residents more senior to the learner. Each EPA includes a 5-point entrustability rating scale and two sections for narrative feedback: 1) "Comments," where assessors provide feedback on what the trainee is excelling at, and 2) "Next Steps," where assessors offer recommendations for further development. Our analysis focused on the quality of feedback in these two sections. At the time, the curriculum comprised of 24 unique EPAs, a target of 107 total EPA assessments, over 24 months of training. The Geriatric Medicine subspecialty training program at the University of Toronto averages four PGY-4 and four PGY-5 subspecialty residents per year with 30 full-time faculty members.

Study Population

The study analyzed all EPAs completed between July 2019 and June 2020, and between July 2021 and June 2022. While the total number of residents in the training program remained constant across both periods, the structure of EPA participation differed. In 2019–2020, only PGY-4 residents engaged in EPA assessments, as PGY-5 residents continued under the previous objective-based curriculum. Consequently, EPA requests were limited to four PGY-4 residents during this period. By 2021–2022, both PGY-4 and PGY-5 residents were integrated into the CBD curriculum, resulting in a total of eight residents actively participating in EPA assessments.

Feedback Analysis

A recently developed qualitative framework was used to evaluate the quality of the narrative feedback. This framework, created by Madrazo *et al*,[15] identified four key domains for high-quality feedback: timeliness, task orientation, actionability, and polarity.

Feedback was considered timely if the EPA was completed within three days of the clinical encounter. The interval between the clinical encounter and the EPA initiation by the learner is termed "time from encounter to trigger" (TET), while the interval from the trigger date to the assessor's completion of the EPA is termed "trigger to completion" (TTC). Timeliness was calculated as the sum of TET and TTC. The three-day benchmark for timeliness was adopted from studies by Hanmore *et al*. and Williams *et al.*,[17,18] and differed from Madrazo and colleagues'[15] definition.

Feedback was labeled as "task-oriented" if it referenced specific tasks or actions. Feedback was categorized as "very actionable" if it included specific recommendations for improvement, "not actionable" if no developmental suggestions were given, and "semi-actionable" for feedback that fell between these extremes.

Narrative feedback was classified as "reinforcing" if it praised the learner's performance, "corrective" if it highlighted areas needing improvement, "mixed" if it included both positive and negative comments, and "neutral" if it either offered no feedback or did not address the learner's performance. Only feedback in the "Next Steps" section was analyzed for polarity, as the "Comments" section was intended to highlight the residents' strengths, differing from Madrazo's *et al*.[15] protocol. Refer to Table 1 for a summary of these definitions and examples.

All identifying information of trainees and assessors was anonymized by the program administrator prior to the study. One investigator (A.D.) reviewed the narratives within each EPA and assigned codes for each quality

domain—timeliness (yes or no), task orientation (yes or no), actionability (very, semi-, or not actionable), and polarity (reinforcing, corrective, mixed, or neutral). A random sample of 100 EPAs was reviewed by a second investigator (A.N.) for independent coding. Coding was performed using Microsoft Excel, and disagreements in coding were resolved through consensus discussions.

Statistical Analysis

Statistical analyses were completed using SAS, version 9.4 (SAS Institute, Inc., Cary, NC). Interrater reliability for each domain within the feedback analysis was determined using Cohen's kappa (for nominal/binary variables) or weighted kappa (for ordinal variables). The level of agreement was interpreted as no (≤0.20), minimal (0.21–0.39), weak (0.40–0.59), moderate (0.60–0.79), strong (0.80–0.90), or almost perfect (>0.90).[19]

Chi-square or Fisher's exact tests were used to compare the type of feedback provided to compare timely versus not timely feedback, direct versus indirect observations, and between the two timeframes.

This study was approved by Sunnybrook Health Sciences Centre Research and Ethics Board.

RESULTS

Overview of 2019–2020 Data

In the 2019–2020 academic year, PGY-4 residents initiated a total of 404 EPA assessments which were completed by assessors and included in our analysis.

Analysis of the feedback showed that 69% (278 of 404) of EPA assessments were timely. Median time for TET was 0 days (25th and 75th percentiles: 0 and 0 days). Median time for TTC was 0 days (25th and 75th percentiles: 0 and 3 days). Eighty-nine percent (360 of 404) of feedback was task oriented. With regard to actionability, 32% (128 of 404) was not actionable, 18% (73 of 404) was semi-actionable, and 50% (203 of 404) was very actionable. Regarding the polarity of feedback, 3% (11 of 404) was reinforcing, 7% (27 of 404) was mixed, 62% (250 of 404) was corrective, and 29% (116 of 404) was neutral.

Overview of 2021-2022 Data

In the 2021–2022 academic year, PGY-4 and -5 residents initiated 851 EPAs. Of these, 645 EPAs (76%) were completed by assessors and included in our analysis.

Analysis of the feedback showed that 64% (410 of 645) of EPAs were timely. Median time for TET was one day (25th

TABLE 1.
Domains of entrustable professional activity feedback analysis

Categories		*Definition*	*Examples*
Timely	*Yes*	*≤ 3 days*	
	No	*> 3 days*	
Task-oriented	Yes	Comments pertained to specific tasks or actions rather than learner attributes	"Strong comprehensive assessment with identification of all major risk factors. Proposed management plan very clear and comprehensive. Clear and organized presentation."
	No		"Great assessment!"
Actionable	Not actionable	No recommendations given	"None identified."
	Semi-actionable	Recommendations that did not name a specific action or behaviour but might identify a more general task or skillset to improve upon	"Continue reading about geriatrics."
	Very actionable	Recommendations that target specific actions or behaviors	"Read about diagnostic criteria for NPH."
Polarity[a]	Reinforcing	Comments endorsed or complimented the residents' performance	"Well done."
	Corrective	Comments indicated problematic performance or need for improvement	"Incorporate driving safety into management of cognitive impairment patients early on."
	Mixed	Comments contained both positive and negative elements	"Management recommendations by resident were holistic and evidence-based. Resident should review clinical clues that help to differentiate vascular or mixed etiology."
	Neutral	Comments were left blank or described the clinical context without providing any comment on resident performance	"None identified."

Adapted from Madrazo *et al*.[15]
[a]Polarity was only assessed within the "Next Steps" section of the EPAs and did not take into consideration the "Comments" section.
NPH = normal pressure hydrocephalus.

and 75th percentiles: 0 and 3 days). Median time for TTC was one day (25th and 75th percentiles: 0 and 2 days). Seventy-eight percent (501 of 645) of feedback was task oriented. With regard to actionability, 47% (305 of 645) was not actionable, 13% (81 of 645) was semi-actionable, and 40% (259 of 645) was very actionable. Regarding the polarity of feedback, 4% (24 of 645) was reinforcing, 5% (35 of 645) was mixed, 47% (305 of 645) was corrective, and 44% (281 of 645) was neutral.

Interrater reliability of investigators was perfect for timeliness (κ=1.000) and almost perfect for task orientation (κ=0.988), actionability (κ=0.988), and polarity (κ=0.994).

Differences Between Timely and Not Timely Feedback

Table 2 presents the type of feedback provided, stratified by timeliness. The chi-square test showed that in the first timeframe, timely feedback was corrective in polarity ($\chi^2(3)=12.64$, $p=.006$). In the second timeframe, timely feedback was task-oriented ($\chi^2(1)=11.87$, $p<.001$), actionable ($\chi^2(2)=24.40$, $p<.001$), and corrective in polarity ($\chi^2(3)=22.80$, $p<.001$). No other differences were identified between timely and not timely feedback in the first timeframe.

Differences Between Feedback Quality Based on Overall Rating

In both timeframes, over 90% of EPAs were given an overall rating of entrustment, with less than 10% requiring minor redirection. Due to the limited number of EPAs needing redirection, further statistical analysis based on the overall rating was not conducted.

Differences Between Feedback Quality Based on Type of Observation

Table 3 presents the type of feedback based on whether the supervisor observed the resident directly or indirectly. Chi-square and Fisher's exact tests were used. Feedback that was given based on indirect observation was more likely to be non-actionable in the first timeframe ($\chi^2(4)=12.34$, $p=.015$)

TABLE 2.
Type of feedback stratified by timely versus not timely

Feedback Domains	*2019–2020*			*2021–2022*		
	Not Timely (n=126), n (%)	*Timely (n=278), n (%)*	*χ^2 (p value)*	*Not Timely (n=235), n (%)*	*Timely (n=410), n (%)*	*χ^2 (p value)*
Task oriented	111 (88)	249 (90)	0.19 (.660)	165 (70)	336 (82)	11.87 (<.001)
Actionable			4.07 (.131)			24.10 (<.001)
Not actionable	45 (36)	83 (30)		140 (60)	165 (40)	
Semi-actionable	27 (21)	46 (17)		28 (12)	53 (13)	
Very actionable	54 (43)	149 (54)		67 (29)	192 (47)	
Polarity			12.64 (.006)			22.80 (<.001)
Reinforcing	6 (5)	5 (2)		10 (4)	14 (3)	
Mixed	15 (12)	12 (4)		11 (5)	24 (6)	
Corrective	67 (53)	183 (66)		84 (36)	221 (54)	
Neutral	38 (30)	78 (28)		130 (55)	151 (37)	

TABLE 3.
Type of feedback stratified by type of observation

Feedback Domains	*2019–2020*			*2021–2022*		
	Direct Observation (n=157), n (%)	*Indirect Observation (n=230), n (%)*	*χ^2 (p value)*	*Direct Observation (n=263), n (%)*	*Indirect Observation (n=342), n (%)*	*χ^2 (p value)*
Timeliness	108 (69)	160 (70)	0.85 (.653)	178 (68)	206 (60)	3.63 (.163)
Task oriented	139 (89)	205 (89)	0.49 (.782)	208 (79)	265 (78)	0.29 (.866)
Actionable			12.34 (.015)			6.34 (.175)
Not actionable	40 (26)	80 (35)		115 (44)	171 (50)	
Semi-actionable	39 (25)	34 (15)		34 (13)	45 (13)	
Very actionable	78 (50)	116 (50)		114 (43)	126 (37)	
Polarity			(.167)[a]			14.93 (.021)
Reinforcing	4 (3)	6 (3)		16 (6)	7 (2)	
Mixed	15 (10)	12 (5)		17 (7)	14 (4)	
Corrective	102 (65)	139 (60)		131 (50)	157 (46)	
Neutral	36 (23)	73 (32)		99 (38)	164 (48)	

[a]Fisher's exact test.

and neutral in polarity in the second timeframe ($\chi^2(6)=14.93$, $p=.021$). There were no other differences in the quality of feedback provided for direct and indirect observation in the first or second timeframe.

Differences Between the Two Timeframes

Table 4 presents the types of feedback provided, stratified by academic year. Analyses were conducted using chi-square tests. Compared to the second timeframe, EPAs completed in the first timeframe were more likely to be task oriented ($\chi^2(1)=22.08$, $p<.001$), actionable ($\chi^2(2)=25.54$, $p<.001$), and corrective in polarity ($\chi^2(3)=25.89$, $p<.001$). No difference in timeliness ($\chi^2(1)=3.03$, $p=.082$) was identified between the feedback provided in the two timeframes.

DISCUSSION

Overall, based on Madrazo *et al*.[15] adapted framework, the quality of narrative feedback delivered to subspecialty geriatric medicine residents at the University of Toronto over both analyzed timeframes was high. The timely feedback was consistently task-oriented, actionable, and corrective in polarity. Significantly and aligned with the true spirit of CBME, feedback provided after direct observation was higher quality, with fewer EPAs being non-actionable or neutral in polarity. The high quality feedback in our program can be partly attributed to our small program size relative to core internal medicine training programs.[16] Other factors may include multiple faculty development sessions on delivering feedback before and during CBD implementation, educational tips on feedback provided within other divisional communication such as newsletters, and the overall culture of feedback within the Department of Medicine.

We analyzed EPAs from two timeframes: 2019–2020, the year CBME was implemented in the geriatric medicine program, and 2021–2022, two years post-implementation. Although there were no differences in timeliness across these periods, EPAs from the initial timeframe were more task-oriented, actionable, and corrective in polarity, suggesting that despite increased familiarity with the curriculum, feedback quality was not sustained over time. Assessment fatigue, the immediate post-COVID-19 pandemic context, lack of real-time quality monitoring, are all theorized as possible explanations.

Similarly, Del Fernandes *et al*.[20] examined the quality of EPAs over a five-year period in the surgical foundations program at Queen's University. They found no clinically significant changes in the quality of narrative feedback over time. Although these findings contrast with ours, as they did not observe a decline in feedback quality, the difference may be attributed to the earlier implementation of CBME in their program, as suggested by previous research.[16] Nonetheless, Del Fernandes *et al*.[20] also identified physician time constraints, excessive administrative burden, and inadequate faculty development as potential contributors to the lack of improvement in EPA quality, which aligns with some of the factors we proposed in our study. Despite the high-quality feedback provided to geriatric residents in our program, there is still room for improvement, specifically in the actionable and corrective polarity domains. Potential methods to improve and sustain feedback quality may include ensuring sustained timeliness, integrating EPA feedback quality reminders within the forms themselves, providing individual routine performance reports to assessors regarding the quality of their feedback, and offering ongoing training sessions.[21,22]

Our study had several limitations. First, it was conducted within a single program which may limit the generalizability of our findings to other geriatric subspecialty training programs. Written feedback recorded in EPAs may not fully capture the verbal feedback provided during clinical interactions, which are occasionally viewed as more effective and honest.[5] Moreover, our methodology did not include the learner's or assessor's perspectives on the feedback they received and delivered, respectively. That type of qualitative evidence may partially complete the gap between in the moment verbal feedback and the written feedback analyzed.

TABLE 4.
Type of feedback stratified by academic year

Feedback Domains	*2019–2020 (n=404), n (%)*	*2021–2022 (n=645), n (%)*	*χ^2 (P value)*
Timeliness	278 (69)	410 (64)	3.03 (.082)
Task oriented	360 (89)	501 (78)	22.08 (<.001)
Actionable			25.54 (<.001)
Not actionable	128 (32)	305 (47)	
Semi-actionable	73 (18)	81 (13)	
Very actionable	203 (50)	259 (40)	
Polarity			25. 89 (<.001)
Reinforcing	11 (3)	24 (4)	
Mixed	27 (7)	35 (5)	
Corrective	250 (62)	305 (47)	
Neutral	116 (29)	281 (44)	

We did not investigate reasons behind low-quality feedback forms which could be due to time constraints, assessment fatigue, electronic system limitations, deficiencies in feedback provider skills, or administrative challenges.

Future studies should explore more detailed factors contributing to the quality of EPA feedback, along with interviews with residents and assessors to better understand their experiences with CBME- and EPAs-based feedback. Additionally, future research should validate a universally agreed-upon scale for assessing the quality of feedback forms.

CONCLUSION

Our study demonstrates that EPAs in a geriatric medicine training program can yield high-quality, actionable feedback when based on direct observation and delivered in a timely manner. The lack of improvement in feedback quality over time highlights the need for ongoing feedback quality assessment and investigations into root causes.

ACKNOWLEDGEMENTS

The authors extend their gratitude to Elizabeth Cho for her assistance in extracting and anonymizing the assessments, Cindy Pham for her assistance with the administrative process, Emma Wilson for her support in conducting the literature search, Dr. Lotus Alphonsus for her guidance on data analysis, Dr. Xingshan Cao for his assistance with statistical analysis, and Drs. Jenny Thain and Sheri-Lynn Kane for their collaboration. We also thank the administrative team members, clinical faculty on the geriatric medicine units, and trainees for their invaluable time and effort in contributing to the assessments used in this study.

CONFLICT OF INTEREST DISCLOSURES

We have read and understood the *Canadian Geriatrics Journal*'s policy on disclosing conflicts of interest and declare that we have none.

FUNDING

This work was supported by the Summer Older Adult Research (SOAR) Experienced Award, funded by private donors to support medical student interest in scholarly activity focused on care of older adults. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, *et al*. Competency-based medical education: theory to practice. *Med Teach*. 2010 Aug 1;32(8):638–45. doi:10.3109/0142159X.2010.501190
2. Lockyer J, Carraccio C, Chan MK, Hart D, Smee S, Touchie C, *et al*. Core principles of assessment in competency-based medical education. *Med Teach*. 2017 Jun 3;39(6):609–16. doi:10.1080/0142159X.2017.1315082
3. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach*. 2007 Jan 1;29(9-10):855–71. doi:10.1080/01421590701775453
4. ten Cate O, Chen HC, Hoff RG, Peters H, Bok H, van der Schaaf M. Curriculum development for the workplace using entrustable professional activities (EPAs): AMEE Guide No. 99. *Med Teach*. 2015 Nov 2;37(11):983–1002. doi:10.3109/0142159X.2015.1060308
5. Tomiak A, Braund H, Egan R, Dalgarno N, Emack J, Reid MA, *et al*. Exploring how the new entrustable professional activity assessment tools affect the quality of feedback given to medical oncology residents. *J Cancer Educ*. 2020 Feb;35(1):165–77. doi:10.1007/s13187-018-1456-z
6. Day BL, Miles A, Ginsburg S, Melvin L. Resident perceptions of assessment and feedback in competency-based medical education: a focus group study of one internal medicine residency program. *Acad Med*. 2020 Feb;95(11):1712–17. doi:10.1097/ACM.0000000000003315
7. Ginsburg S, van der Vleuten CP, Eva KW, Lingard L. Cracking the code: residents' interpretations of written assessment comments. *Med Educ*. 2017 Apr;51(4):401–10. doi:10.1111/medu.13158
8. Holmboe ES, Yamazaki K, Edgar L. Reflections on the first 2 years of milestone implementation. *J Grad Med Educ*. 2015;7(3):506–11. doi:10.4300/JGME-07-03-43
9. Lefroy J, Watling C, Teunissen PW, Brand P. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. *Perspect Med Educ*. 2015;4(6):284–99. doi:10.1007/s40037-015-0231-7
10. Weston PS, Smith CA. The use of mini-CEX in UK foundation training six years following its introduction: lessons still to be learned and the benefit of formal teaching regarding its utility. *Med Teach*. 2014 Feb 1;36(2):155–63. doi:10.3109/0142159X.2013.836267
11. Dudek NL, Marks MB, Wood TJ, Lee AC. Assessing the quality of supervisors' completed clinical evaluation reports. *Med Educ*. 2008 Aug;42(8):816–22. doi:10.1111/j.1365-2923.2008.03105.x
12. Young JQ, Sugarman R, Holmboe E, O'Sullivan PS. Advancing our understanding of narrative comments generated by direct observation tools: lessons from the psychopharmacotherapy-structured clinical observation. *J Grad Med Educ*. 2019 Oct 1;11(5):570–79. doi:10.4300/JGME-D-19-00207.1
13. Bentley H, Darras KE, Forster BB, Sedlic A, Hague CJ. Review of challenges to the implementation of competence by design in post-graduate medical education: what can diagnostic radiology learn from the experience of other specialty disciplines? *Acad Radiol*. 2022 Dec 1;29(12):1887–96. doi:10.1016/j.acra.2021.11.025
14. Shafian S, Ilaghi M, Shahsavani Y, Okhovati M, Soltanizadeh A, Aflatoonian S, *et al*. The feedback dilemma in medical education: insights from medical residents' perspectives. *BMC Med Educ*. 2024 Apr 19;24(1):424. doi:10.1186/s12909-024-05398-y
15. Madrazo L, DCruz J, Correa N, Puka K, Kane SL. Evaluating the quality of written feedback within entrustable professional activities in an internal medicine cohort. *J Grad Med Educ*. 2023 Feb 1;15(1):74–80. doi:10.4300/JGME-D-22-00222.1
16. Clement EA, Oswald A, Ghosh S, Hamza DM. Exploring the quality of feedback in entrustable professional activity narratives across 24 residency training programs. *J Grad Med Educ*. 2024 Feb 1;16(1):23–29. doi:10.4300/JGME-D-23-00210.1

17. Hanmore T, Moon CC, Curtis R, Hopman W, Baxter S. Is time really of the essence? Timeliness of narrative feedback in ophthalmology CBME assessments. *Med Teach*. 2024 May 3;46(5):705–10. doi:10.1080/0142159X.2023.2274286
18. Williams RG, Chen X (Phoenix), Sanfey H, Markwell SJ, Mellinger JD, Dunnington GL. The measured effect of delay in completing operative performance ratings on clarity and detail of ratings assigned. *J Surg Educ*. 2014 Nov-Dec;71(6):e132–e138. doi:10.1016/j.jsurg.2014.06.015
19. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Medica*. 2012 Oct 15;22(3):276–82. doi:10.11613/BM.2012.031
20. Del Fernandes R, de Vries I, McEwen L, Mann S, Phillips T, Zevin B. Evaluating the quality of narrative feedback for entrustable professional activities in a surgery residency program. *Ann Surg*. 2024 Dec 1;280(6):916–24. doi:10.1097/SLA.0000000000006308
21. Faiella W, Moeller A. Competency-based cardiology training: a simple approach to improve supervisor completion of EPAs. *Can J Cardiol*. 2022 Oct;38(10 Suppl 2):S134–S135. doi:10.1016/j.cjca.2022.08.069
22. Renting N, Jaarsma D, Borleffs JC, Slaets JP, Cohen-Schotanus J, Gans RO. Effectiveness of a supervisor training on quality of feedback to internal medicine residents: a controlled longitudinal multicentre study. *BMJ Open*. 2023 Sep 1;13(9):e076946. doi:10.1136/bmjopen-2023-076946

Correspondence to: Dov Gandell, MDCM, Sunnybrook Health Sciences Centre, 2075 Bayview Ave., Toronto, ON M4N 3M5
E-mail: dov.gandell@sunnybrook.ca