

Original Article

Self-assessment of Competence in Endoscopy: Challenges and Insights

Michael A. Scaffidi, BSc, MEd^{1,2,✉}, Rishad Khan, BSc, MD^{2,3}, Samir C. Grover, MD, MEd, FRCPC, AGAF^{2,4}, Nikko Gimpaya, BSc², Catharine M. Walsh, MD, MEd, PhD, FAAP, FRCPC^{5,6,7}

¹School of Medicine, Faculty of Health Sciences, Queen's University, Kingston, Canada; ²Division of Gastroenterology, St. Michael's Hospital, University of Toronto, Toronto, Canada; ³Department of Medicine, Schulich School of Medicine and Dentistry, Western University, London, Canada; ⁴Department of Medicine, University of Toronto, Toronto, Canada; ⁵Division of Gastroenterology, Hepatology, and Nutrition and the Research and Learning Institutes, Hospital for Sick Children, University of Toronto, Toronto, Canada; ⁶Department of Paediatrics, Faculty of Medicine, University of Toronto, Toronto, Canada; ⁷The Wilson Centre, Faculty of Medicine, University of Toronto, Toronto, Canada

Correspondence: Catharine M. Walsh, MD, MEd, PhD, FAAP, FRCPC, Division of Gastroenterology, Hepatology and Nutrition, the Learning and Research Institutes, Hospital for Sick Children, Department of Paediatrics and the Wilson Centre, Faculty of Medicine, University of Toronto, SickKids PGCRL, 686 Bay Street, Room 11.9719, Toronto, Ontario M5G 0A4, Canada, e-mail: catharine.walsh@utoronto.ca

Abstract

Background: Endoscopists use self-assessment to monitor the development and maintenance of their skills. The accuracy of these self-assessments, which reflects how closely one's own rating corresponds to an external rating, is unclear.

Methods: In this narrative review, we critically examine the current literature on self-assessment in gastrointestinal endoscopy with the aim of informing training and practice and identifying opportunities to improve the methodological rigor of future studies.

Results: In the seven included studies, the evidence regarding self-assessment accuracy was mixed. When stratified by experience level, however, novice endoscopists were least accurate in their self-assessments and tended to overestimate their performance. Studies examining the utility of video-based interventions using observation of expert benchmark performances show promise as a mechanism to improve self-assessment accuracy among novices.

Conclusions: Based on the results of this review, we highlight problematic areas, identify opportunities to improve the methodological rigor of future studies on endoscopic self-assessment and outline potential avenues for further exploration.

Keywords: *Clinical competence/standards; Educational measurement; Endoscopy; Gastrointestinal/education; Gastrointestinal/standards; Self-assessment*

Introduction

Assessment guides the development and maintenance of competence in medicine (1). Although gastrointestinal endoscopy trainees are commonly assessed by an external rater (2,3), self-assessment has been proposed as a means by which complementary data can be collected (4). Self-assessment, defined as “a personal evaluation of one's professional attributes and abilities against

perceived norms” (5), has implications as a meta-cognitive skill that underpins continual skill development (6–8). In particular, self-assessment has a role in the management of one's own learning (9), as one study demonstrated that improved self-assessment skills led to more fruitful self-regulated learning (10). Given the impact of self-regulated learning in clinical contexts (11), the application of self-assessment requires careful consideration.

Specifically, self-assessment must be accurate—that is, one’s self-assigned rating should correspond to one’s skill level and align with externally assigned ratings (12). Current evidence, however, strongly suggests that health professionals cannot accurately assess their own performances with respect to clinical skills (e.g., diagnostic ability) and cognitive skills (e.g., general medical knowledge) (13–15). Conversely, there is uncertainty about whether self-assessment is more accurate for procedural skills. In particular, a narrative review of surgeons found mixed results regarding procedural self-assessment accuracy (16).

In gastrointestinal endoscopy, the importance of self-assessment has been acknowledged by credentialing bodies, such as the Joint Advisory Group on Gastrointestinal Endoscopy (JAG), which recommends that trainees incorporate self-assessment practices into their training (17). Additionally, it has been emphasized as an important component of maintenance of certification by professional organizations, such as the Royal College of Physicians and Surgeons of Canada, to enable identification of opportunities to enhance competence through future learning activities (18). To date, however, the evidence surrounding endoscopic self-assessment has yet to be thoroughly examined. Although there are systematic reviews of assessment in endoscopy (3,19), self-assessment was addressed only briefly. In this narrative review, we provide a summative, critical examination of the endoscopic self-assessment literature.

Current State of Evidence of Endoscopic Self-Assessment

There is a growing evidence base on endoscopic self-assessment as we identified seven studies on the topic (see Table 1) (20–26). One of the first studies that presented data on endoscopic self-assessment accuracy was published by Vassiliou et al. in 2010 (20). Although the primary focus of the paper was the validation of two technical skills assessment tools, the Global Assessment of Gastrointestinal Endoscopic Skills for upper endoscopy (GAGES-UE) and for colonoscopy (GAGES-C), self-assessment accuracy data was presented as a secondary outcome (20). The authors reported good agreement between external and self-assessed scores using both tools. Another study examined self-assessment in the context of simulated colonoscopic polypectomy (21). The authors reported an overall weak correlation between externally and self-assessed scores using the Direct Observation of Polypectomy Skills (DOPyS) tool, among a group that included endoscopists of novice, intermediate, advanced and expert experience levels (27). When the analysis was stratified by experience level, advanced and expert endoscopists were shown to have relatively more accurate self-assessment, albeit they were still considered “inaccurate.” Of note, the authors found that novices and intermediates tended to underestimate their performances, while advanced and expert endoscopists tended toward overestimation. The

Table 1. Summary of the self-assessment literature in gastrointestinal endoscopy

First author and year	Methodology			Findings		
	Endoscopic setting	Endoscopist sample	Endoscopic procedure(s)	Overall accuracy of self-assessment	Direction of self-assessment bias (if present)	Effect of endoscopic experience on self-assessment accuracy
Vassilou (2010)	Clinical	Adult endoscopists	Colonoscopy, gastroscopy	Accurate	N/A	N/A
Ansell (2014)	Simulated	Adult	Colonoscopy	Inaccurate	Overestimation	Improved
Moritz (2016)	Simulated	Adult	Colonoscopy	Inaccurate	Underestimation	No change
Vyasa (2016)	Simulated	Adult	Sigmoidoscopy	Accurate	Overestimation	N/A
Scaffidi (2018a)	Clinical	Adult	Colonoscopy	Accurate	Overestimation	Improved
Scaffidi (2018b)	Clinical	Pediatric	Colonoscopy	Accurate	Overestimation	Improved
Scaffidi (2019)	Simulated	Adult	Gastroscopy	Accurate	Overestimation	N/A

N/A, not applicable

two key findings from this study, inaccurate overall self-assessment and underestimation of performance, were also found in the study by Moritz et al. on self-assessed colonoscopy quality metrics (22). In particular, the primary analysis revealed that endoscopists were inaccurate as they underestimated their performance on several parameters of colonoscopy quality that were recorded in a quality assurance registry, including cecal intubation rate and withdrawal time. Furthermore, there were no significant differences when accounting for endoscopic experience or for gender, although the authors remark that their study may have been underpowered to detect these differences.

While the previous studies suggest a minimal to negligible effect of endoscopic experience on self-assessment accuracy, two recent studies indicate otherwise. A 2018 study by Scaffidi et al. that aimed to evaluate self-assessment accuracy of novice, intermediate and experienced endoscopists found that there was a positive relationship between experience level and self-assessment accuracy of clinical competence (24). Although there was an overall moderate agreement between externally and self-assessed scores on the Gastrointestinal Endoscopic Colonoscopy Assessment Tool (GiECAT), a procedure-specific assessment tool, experienced endoscopists were significantly more accurate in their self-assessments compared with novices (28). Furthermore, experienced endoscopists tended to underestimate their performances, while novices tended to overestimate themselves. A subsequent study by the same group in the setting of pediatric clinical colonoscopies yielded similar results (25). Specifically, compared with novices, experienced endoscopists had more accurate self-assessments using the Gastrointestinal Endoscopy Competency Assessment Tool for Pediatric Colonoscopy (GiECAT_{KIDS}), wherein novices overestimated their performances and experienced endoscopists underestimated (29). Taken together, these studies provide compelling evidence that experience can positively influence self-assessment accuracy. This finding has a strong theoretical basis, which will be discussed in a later section.

A consistent trend in the extant literature is that novice endoscopists demonstrate inaccurate self-assessment, a deficiency that two studies sought to remedy using video-based feedback (23,26). The first, a trial published by Parth et al. in 2016, involved general surgery residents self-assessing their ability to perform a simulated screening colonoscopy using procedural metrics (e.g., time to cecum) generated by a virtual reality simulator (23). After completing a pretest, participants were randomized into one of three arms: video review of their own simulated procedures; video review of an expert completing the same procedure; or additional practice on the simulator (instead of video review) as a control. Participants then completed 10 additional simulated practice cases after the intervention, followed by a posttest on the same initial simulated case. The authors reported a small improvement in self-assessment accuracy among participants who watched the expert

benchmark video. The second, a study by Scaffidi et al., used a similar methodology, although the simulated procedure was esophagoduodenoscopy (EGD) and participants, while all novice endoscopists, were recruited across different specialties and varying levels of training, which included medical students, residents and staff physicians (26). After completing a simulated EGD procedure as a pretest, participants were randomized to one of three interventions: video review of their own simulated EGD case; video review of a benchmark (i.e., expert) performance of the simulated EGD case; or access to both videos for review. Participants then completed two posttests: the same simulated EGD case as the pretest, which determined skill retention, and a novel EGD case, which determined transfer. Self-assessments were completed during all testing using the global rating scale component of the GiECAT (28). The authors found that both the benchmark and combination interventions led to improved self-assessment accuracy, although the former group only demonstrated improvement on the retention test, while the latter group only on the transfer test, suggesting that combination video review leads to more sustained improvements in self-assessment accuracy.

Challenges with Endoscopic Self-Assessment

Despite the growing evidence base regarding endoscopic self-assessment, a summative interpretation of existing studies is limited by several key methodological deficiencies. The first pertains to the assessment metrics utilized. Studies aiming to investigate self-assessment should ensure that the assessment metrics employed have adequate validity evidence. Much of the current literature utilized simulator metrics (23), quality metrics and/or procedural endpoints (e.g., procedure duration) (22) as surrogate measures of endoscopic competence. Although these are easy to calculate and interpret, there is mixed evidence regarding their validity as markers of endoscopic competence (3,19). Moreover, these often singular-measure approaches fail to holistically capture the multifaceted dimensions of endoscopic competence, including those that relate to cognitive (e.g., pathology recognition), technical (e.g., scope navigation) and integrative (e.g., communication) skills (2). There are a number of direct observational procedural assessment tools with strong validity evidence that have been recommended for assessment of endoscopic competence (2,19), such as the DOPyS (21), the GiECAT (24,26), or the GiECAT_{KIDS} (25), which have been used in the more recent studies of higher methodological quality.

Second, there are several statistical techniques that have been used inappropriately in previous studies of endoscopic self-assessment. In particular, linear correlation measures, such as Spearman's rho (21), cannot be used to capture agreement between external- and self-assessors because they are prone to systematic variations in the data, such as chance

agreement between assessors (16,30). Another, suboptimal method used is the scatterplot of external and self-assessed scores (22), which, while not necessarily incorrect, fails to adequately detect systematic subtleties (i.e., overestimation and underestimation) between the two sets of scores. Based on the method comparison literature (31), we recommend the following three techniques: Bland–Altman plots, which can visualize systematic differences, such as overestimation or underestimation, by plotting the mean differences between the two sets of raters against the mean values of the scores: the intraclass correlation coefficient (ICC), which quantifies agreement between two or more raters while correcting for chance agreement, and absolute difference scores, which involve the calculation of the absolute differences between external- and self-assessors that can be used to determine if there are differences between groups of endoscopists (e.g., varying experience levels). These techniques were utilized in three of the aforementioned studies (24–26).

Insights Into Endoscopic Self-Assessment

To better understand the findings from the endoscopic self-assessment literature, we consider two relevant concepts from the psychological literature. The first is the Dunning–Kruger effect, which is a well-described psychological phenomenon, wherein those who are least competent overestimate their skill level due to their lack of awareness (32). In terms of endoscopic self-assessment, this phenomenon explains why most of the studies we identified demonstrate that novice endoscopists are the least accurate self-assessors (21,23–26) and why several studies found that novices tend to overestimate their competence (23–26). Furthermore, the Dunning–Kruger effect may provide a mechanistic explanation for the effectiveness of benchmark video-based interventions in improving self-assessment accuracy among novices. Accurate self-assessment requires insight into one's own performance, an awareness of the performance of others and a capacity to reflect on both measures and make a judgment (33). Observation of experts may provide novices with realistic standards for assessing their own performance despite their inexperience (26).

Another possible explanation for inaccurate self-assessment is impression management, whereby individuals intentionally present a more favorable view of themselves (34). A study of self-assessment in oral surgery found that surgeons who inaccurately overestimated their performance also had higher impression management scores, indicating that they had a greater tendency to deliberately distort responses to present a more positive image of themselves (35). This framework has been used to explain why endoscopists, especially novices, may have inaccurate self-assessments that typically involve overestimation as they intend to create a favorable perception of their performances (21).

Implications for Endoscopic Training, Practice and Future Studies

Self-assessment is important for several reasons. It enables the identification of one's weakness to help set appropriate learning goals, enable one to self-limit in areas outside of one's scope of competence practice and to set realistic expectations for oneself. Finally, self-assessment is an integral component in the process of self-regulated learning, wherein one's own learning is guided by meta-cognition (i.e., awareness and understanding of one's own thought processes). Self-regulated learning has important implications for physicians both in training and in practice as skill development continues throughout one's entire career. Research has highlighted the importance of self-assessment in relation to self-regulated learning. A recent paper summarizing four meta-analyses (19 studies) found that self-assessment practices had a considerable impact on individuals' self-regulated learning, with effect sizes ranging from 0.23 to 0.65 (36). Furthermore, the model of self-regulated learning proposed by Winne and Hadwin asserts that self-assessment is essential in that it enables learners to ensure that they are on target with their learning objectives (37).

In general, we found mixed evidence regarding the accuracy of endoscopic self-assessment, with three studies reporting an overall moderate accuracy for endoscopists (20,24,25) and two studies reporting inaccurate self-assessment (21,22). When stratified by endoscopic experience, however, the primary at-risk group for inaccurate self-assessment are novices as they not only lack experience but are unaware of the extent to which they are unskilled (32). Although novice endoscopists demonstrate a clear inability to accurately self-assess, existing evidence is less clear regarding experienced endoscopists. Specifically, three studies demonstrated that more experienced endoscopists had better self-assessment accuracy as compared with inexperienced individuals (21,24,25). It is unclear, however, what threshold of endoscopic experience confers this capacity for improved self-assessment. While two of the aforementioned studies found that experienced endoscopists, who completed more than 1000 colonoscopies, were most accurate in their self-assessments, there was no clear trend in self-assessment accuracy for intermediate endoscopists, who completed between 50 and 500 colonoscopies (24,25).

In terms of endoscopic practice and training, we note several implications. First, endoscopists of any experience level should avoid the use of quality indicators and/or simulator-based metrics as the basis for self-assessment as these lack strong evidence of validity for endoscopic competence (3,19). Additionally, such singular measures provide little informative feedback to help pinpoint deficiencies (2). Instead, tools that focus on direct observation of skills (e.g., DOPyS and GiECAT) should be implemented for self-assessment. One way to implement self-assessment of endoscopic competence is through the use of

video recordings, which have been shown to have good validity evidence (38). Videos can be recorded during the procedure and assessed at a later time by an external- and/or self-assessor, thus allowing more time for reflection on performance. More experienced endoscopists, who generally demonstrate better self-assessment accuracy, can incorporate self-assessment into their practice to help self-identify their learning needs to create individualized plans for their continuing professional development. They may wish, however, to monitor their initial assessments to ensure that they are accurate and can potentially engage in courses that provide structured assessment and feedback (39). Educators leading endoscopy-related professional development activities can also potentially use self-assessment data to tailor educational content to more closely match participants' learning needs.

Conversely, novice endoscopists may require targeted attention with respect to self-assessment. In particular, video-based interventions can positively impact self-assessment accuracy. Two included studies found that novices who watched a benchmark video of an expert perform the same endoscopic procedure, before their own performance, had improved self-assessment accuracy (23,26). Video interventions targeting self-assessment accuracy have also demonstrated effectiveness in other skill domains, such as surgery (40).

External feedback from an expert can also play a key role in regard to self-assessment accuracy as the views of others can be incorporated to create the self-awareness necessary for learning. Eva and Regehr suggest that self-improvement requires not only reliable external feedback but also deliberate efforts to reflect on that feedback so that it becomes meaningful (41). Additionally, the combination of external feedback and self-assessment appears to be synergistic; a meta-analysis found improved learning when self-assessment was combined with feedback as opposed to self-assessment alone (42). However, literature has shown that the perceived quality of feedback has an impact on the likelihood that individuals will use the feedback to inform their self-assessment and improve their practices (43). It is important for endoscopic trainers to consider the manner in which novices engage in self-assessment and receive feedback. In line with the Dunning–Kruger phenomenon discussed earlier, less experienced individuals lack accurate insight into their performances and, therefore, tend to overestimate their awareness (32). This may also affect novices' perception of feedback. An overconfident novice, for example, may dismiss negative feedback from an external assessor as unreliable. On the other hand, feedback delivered from a clear position of beneficence has been shown to be more likely perceived as worthy of attention (7). To ensure that feedback is understood and acted upon, endoscopic trainers should consider using *informed* self-assessment, wherein they review self-assessment data with trainees to foster self-reflection and help them to develop a more accurate

perception of their abilities and learning deficits (7,44). This will, ultimately, enable them to become more self-aware and nurture their ability to self-direct their own learning.

There is a growing interest in the use of “objective” measures of performance, such as cecal intubation rate. While such quality metrics are attractive, there is evidence indicating that reliance on singular end-point measure may have unintended consequences as endoscopists may “game” certain metrics while disregarding others, which can hinder comprehensive feedback (45). A study by Razzak et al., which examined the impact of monitoring both adenoma detection rate and cecal withdrawal time, found that monitoring only withdrawal time led to fewer adenomas removed per examination (46). It is, therefore, unclear whether this strategy should be used among trainees, as their inexperience may potentially even exacerbate the unintended effect of “gaming” metrics. In practice, it has been suggested that a combination of metrics is preferable to provide a more comprehensive picture of performance (47).

There are opportunities to improve future research in the area of endoscopic self-assessment. First, the assessment measures used should have strong evidence of validity as evaluated through the lens of a validity framework, such as the one described by Messick (48,49). A recent systematic review used Messick's framework to evaluate the validity evidence of methods for certification in colonoscopy (3). Several assessment tools, such as the DOPyS and the GiECAT, used in the more recent, higher-quality studies of endoscopic self-assessment (24–26), have strong evidence of validity. Second, appropriate measurement of self-assessment accuracy should be used. Specifically, we suggest three statistical techniques from the method comparison literature—the ICC; Bland–Altman plots and absolute difference scores—as these are better suited to detect systematic differences in agreement among individuals (31). Finally, moderators of self-assessment, such as experience level (24,25), should be carefully considered in any investigation as the current findings demonstrate that they have an impact on self-assessment accuracy.

In terms of future studies on endoscopic self-assessment, we suggest three areas for exploration. First, a systematic review that quantitatively summarizes endoscopic self-assessment accuracy would provide an objective means by which to determine if endoscopists are accurate. Second, studies could further explore the impact of video interventions on self-assessment accuracy, for example, through the integration of benchmark videos into existing simulation-based training curricula aimed at novice endoscopists (50–55). Finally, future studies should evaluate the theoretical underpinnings of self-assessment accuracy, such as the measurement of impression management, which has been conducted for other procedures (35).

Conclusions

Self-assessment is an important meta-cognitive skill in the development and maintenance of procedural competence. In this narrative review, we examined the extant literature on endoscopic self-assessment through a summative, critical lens to highlight implications for training, practice and research. In general, we found that novice endoscopists are most inaccurate in their self-assessment, while experienced endoscopists are more accurate. Based on the current literature, we make several recommendations for training and practice, which includes use of direct observation assessment tools with strong validity evidence, use of video-based interventions to improve self-assessment accuracy (especially among novices) and incorporation of self-assessment into practice and training.

Funding

C.M.W. holds a Career Development Award from the Canadian Child Health Clinician Scientist Program and an Early Researcher Award from the Ontario Ministry of Research, Innovation and Science. The funders had no role in the design and conduct of the review, decision to publish and preparation, review, or approval of the manuscript.

Acknowledgments

Study design and planning: all authors; drafting of the manuscript: M.A.S. and C.M.W.; critical revision of manuscript for important intellectual content: all authors; approval of final version of manuscript: all authors.

Conflicts of Interest

R.K. has received research funding from AbbVie, Ferring Pharmaceuticals, and Pendopharm. S.C.G. has received research funding from AbbVie and Janssen and personal fees from AbbVie, Takeda and Ferring, and is owner of, and holds shares, Volô Healthcare. All other authors have no conflicts of interest to disclose.

References

- Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387–96.
- Walsh CM. In-training gastrointestinal endoscopy competency assessment tools: Types of tools, validation and impact. *Best Pract Res Clin Gastroenterol*. 2016;30(3):357–74.
- Preisler L, Svendsen MBS, Svendsen LB, Konge L. Methods for certification in colonoscopy—A systematic review. *Scand J Gastroenterol*. 2018;53(3):350–8.
- Redwood C, Winning T, Townsend G. The missing link: Self-assessment and continuing professional development. *Aust Dent J*. 2010;55(1):15–19.
- Colthart I, Bagnall G, Evans A, et al. The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice: BEME Guide no. 10. *Med Teach*. 2008;30(2):124–45.
- Pintrich PR. The role of metacognitive knowledge in learning the role of metacognitive knowledge in learning, teaching, and assessing. *Theor Pract*. 2002;41(4):219–25.
- Sargeant J, Armson H, Chesluk B, et al. The processes and dimensions of informed self-assessment: A conceptual model. *Acad Med*. 2010;85(7):1212–20.
- Duffy FD, Holmboe ES. Self-assessment in lifelong learning and improving performance in practice physician know thyself. *JAMA* 2006;296(9):1137–9.
- Bjork RA, Dunlosky J, Kornell N. Self-regulated learning: beliefs, techniques, and illusions. *Annu Rev Psychol*. 2013;64:417–44.
- Kostons D, van Gog T, Paas F. Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learn Instr*. 2012;22(2):121–32.
- van Houten-Schat MA, Berkhout JJ, van Dijk N, et al. Self-regulated learning in the clinical context: A systematic review. *Med Educ*. 2018;52(10):1008–15.
- Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. *Adv Health Sci Educ Theory Pract*. 2002;7(1):63–80.
- Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA*. 2006;296(9):1094–102.
- Blanch-Hartigan D. Medical students' self-assessment of performance: Results from three meta-analyses. *Patient Educ Couns*. 2011;84(1):3–9.
- Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Acad Med*. 1991;66(12):762–9.
- Zevin B. Self versus external assessment for technical tasks in surgery: A narrative review. *J Grad Med Educ*. 2012;4(4):417–24.
- Joint Advisory Group on Gastrointestinal Endoscopy. Joint Advisory Group on Gastrointestinal Endoscopy (JAG) Accreditation Standards for Endoscopy Services. 2014. <https://www.rcplondon.ac.uk/file/3890/download> (Accessed June 5, 2020).
- Royal College of Physicians and Surgeons of Canada. CPD accreditation: Self-assessment programs. Royal College Web site. <http://www.royalcollege.ca/rcsite/cpd/accreditation/cpd-accreditation-self-assessment-programs-saps-e> (Accessed June 5, 2020).
- Ekkelenkamp VE, Koch AD, de Man R a, Kuipers EJ, Man RA De, Kuipers EJ. Training and competence assessment in GI endoscopy: A systematic review. *Gut*. 2016;65(4):607–15.
- Vassiliou MC, Kaneva PA, Poulouse BK, et al. Global Assessment of Gastrointestinal Endoscopic Skills (GAGES): A valid measurement tool for technical skills in flexible endoscopy. *Surg Endosc*. 2010;24(8):1834–41.
- Ansell J, Hurley JJ, Horwood J, et al. Can endoscopists accurately self-assess performance during simulated colonoscopic polypectomy? A prospective, cross-sectional study. *Am J Surg*. 2014;207(1):32–38.
- Moritz V, Holme O, Leblanc M, et al. An explorative study from the Norwegian Quality Register Gastronet comparing self-estimated versus registered quality in colonoscopy performance. *Endosc Int Open*. 2016;4(3):E326–32.
- Vyasa P, Willis RE, Dunkin BJ, et al. Are general surgery residents accurate assessors of their own flexible endoscopy skills? *J Surg Educ*. 2017;74(1):23–29.
- Scaffidi MA, Grover SC, Carnahan H, et al. Impact of experience on self-assessment accuracy of clinical colonoscopy competence. *Gastrointest Endosc*. 2018;87(3):827–36.e2.
- Scaffidi MA, Khan R, Carnahan H, et al. Can pediatric endoscopists accurately assess their clinical competency? A comparison across skill levels. *J Pediatr Gastroenterol Nutr*. 2019;68(3):311–7.
- Scaffidi MA, Walsh CM, Khan R, et al. Influence of video-based feedback on self-assessment accuracy of endoscopic skills: A randomized controlled trial. *Endosc Int Open*. 2019;7(5):E678–84.
- Gupta S, Anderson J, Bhandari P, et al. Development and validation of a novel method for assessing competency in polypectomy: Direct observation of polypectomy skills. *Gastrointest Endosc*. 2011;73(6):1232–9.e2.
- Walsh CM, Ling SC, Khanna N, et al. Gastrointestinal Endoscopy Competency Assessment Tool: Reliability and validity evidence. *Gastrointest Endosc*. 2015;81(6):1417–24.e2.
- Sauer CG, Narkewicz MR. GiECAT(KIDS) validated pediatric colonoscopy assessment tool: A call to action. *J Pediatr Gastroenterol Nutr*. 2015;60(4):425–7.
- Kim HY. Statistical notes for clinical researchers: Evaluation of measurement error 2: Dahlberg's error, Bland-Altman method, and Kappa coefficient. *Restor Dent Endod*. 2013;38(3):182–5.
- Watson PE, Petrie A. Method agreement analysis: A review of correct methodology. *Theriogenology*. 2010;73(9):1167–79.
- Kruger J, Dunning D. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol*. 1999;77(6):1121–34.
- Hays RB, Jolly BC, Caldon LJ, et al. Is insight important? Measuring capacity to change performance. *Med Educ*. 2002;36(10):965–71.
- Leary MR, Kowalski RM. Impression management: a literature review and two-factor model. *Psychol Bull*. 1990;107(1):34–47.
- Evans AW, Leeson RM, Newton John TR, et al. The influence of self-deception and impression management upon self-assessment in oral surgery. *Br Dent J*. 2005;198(12):765–9; discussion 755.
- Panadero E, Jonsson A, Botella J. Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educ Res Rev*. 2017;22(November):74–98.
- Panadero E. A review of self-regulated learning: six models and four directions for research. *Front Psychol*. 2017;8(April):422.

38. Scaffidi M, Grover SC, Carnahan H, et al. A prospective comparison of live and video-based assessments of colonoscopy performance. *Gastrointest Endosc.* 2017;87(3):766–75.
39. Dubé C, Rostom A. Acquiring and maintaining competency in gastrointestinal endoscopy. *Best Pract Res Clin Gastroenterol.* 2016;30(3):339–47.
40. Hawkins SC, Osborne A, Schofield SJ, et al. Improving the accuracy of self-assessment of practical clinical skills using video feedback—The importance of including benchmarks. *Med Teach.* 2012;34(4):279–84.
41. Eva KW, Regehr G. Self-assessment in the health professions: A reformulation and research agenda. *Acad Med.* 2005;80(10 Suppl):S46–54.
42. Sitzmann T, Ely K, Brown KG, Bauer KN. Self-assessment of knowledge: A cognitive learning or affective measure? Perspectives from the management learning and education community. *Acad Manag Learn Educ.* 2010;9(2):335–41.
43. Mann K, van der Vleuten C, Eva K, et al. Tensions in informed self-assessment: How the desire for feedback and reticence to collect and use it can conflict. *Acad Med.* 2011;86(9):1120–7.
44. Sargeant J, Eva KW, Armson H, et al. Features of assessment learners use to make informed self-assessments of clinical performance. *Med Educ.* 2011;45(6):636–47.
45. Thomas-Gibson S, Haycock A, Valori R. The intended and unintended consequences of performance monitoring in colonoscopy. *Endosc Int Open.* 2016;4(10):E1028–9.
46. Razzak A, Smith D, Zahid M, Papachristou G, Khalid A. Effect of quality metric monitoring and colonoscopy performance. *Endosc Int Open.* 2016;4(10):E1023–7.
47. Valori RM, Damery S, Gavin DR, et al. A new composite measure of colonoscopy: The Performance Indicator of Colonic Intubation (PICI). *Endoscopy.* 2018;50(1):40–51.
48. Messick S. Validity. In: Linn RL, ed. *Educational Measurement*, 3rd ed. New York, NY: American Council on Education and Macmillan; 1989: 13–104.
49. Beckman TJ, Cook DA, Mandrekar JN. What is the validity evidence for assessments of clinical teaching? *J Gen Intern Med.* 2005;20(12):1159–64.
50. Grover SC, Garg A, Scaffidi MA, et al. Impact of a simulation training curriculum on technical and nontechnical skills in colonoscopy: A randomized trial. *Gastrointest Endosc.* 2015;82(6):1072–9.
51. Grover SC, Scaffidi MA, Khan R, et al. Progressive learning in endoscopy simulation training improves clinical performance: A blinded randomized trial. *Gastrointest Endosc.* 2017;86(5):881–9.
52. Walsh CM, Scaffidi MA, Khan R, et al. Non-technical skills curriculum incorporating simulation-based training improves performance in colonoscopy among novice endoscopists: Randomized controlled trial. *Dig Endosc.* 2020. Epub ahead of print. doi: [10.1111/den.13623](https://doi.org/10.1111/den.13623)
53. Scaffidi MA, Khan R, Walsh CM, et al. Protocol for a randomised trial evaluating the effect of applying gamification to simulation-based endoscopy training. *BMJ Open.* 2019;9(2):e024134.
54. Haycock A, Koch AD, Familiari P, et al. Training and transfer of colonoscopy skills: A multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointest Endosc.* 2010;71(2):298–307.
55. Khan R, Plahouras J, Johnston BC, et al. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. *Cochrane Database Syst Rev.* 2018;8:CD008237.