Annals of Medicine

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/iann20

Informing factors and outcomes of self-assessment practices in medical education: a systematic review

Binbin Zheng, Qing He & Junru Lei

To cite this article: Binbin Zheng, Qing He & Junru Lei (2024) Informing factors and outcomes of self-assessment practices in medical education: a systematic review, Annals of Medicine, 56:1, 2421441, DOI: 10.1080/07853890.2024.2421441

To link to this article: https://doi.org/10.1080/07853890.2024.2421441

Elevate Series

ANNALS OF MEDICINE
2024, VOL. 56, NO. 1, 2421441
https://doi.org/10.1080/07853890.2024.2421441

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Informing factors and outcomes of self-assessment practices in medical education: a systematic review

Binbin Zheng[a], Qing He[b] and Junru Lei[b]

[a]Uniformed Services University of the Health Sciences, Bethesda, MD, USA; [b]Bau Institute of Medical & Health Sciences Education, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, Hong Kong SAR

ABSTRACT

Background: Self-assessment (SA) is increasingly recognized not only as an assessment method but also as a vital learning activity that enables learners to identify their strengths and limitations, fostering a crucial skill for lifelong learning. Despite its acknowledged significance, there remains a gap in understanding SA *for* and *as* learning, rather than merely *of* learning. This review, therefore, explores two primary questions: (1) What factors inform the implementation of medical students' SA practices? (2) How does SA practices contribute to various educational outcomes?

Methods: A systematic review was conducted across five databases, employing a combination of keywords pertinent to SA and medical education. Our selection criteria deliberately excluded articles that solely focused on the accuracy of SA or used SA exclusively as a measure for other outcome variables. As a result, 39 studies met our inclusion criteria and were analyzed for this review. Qualitative narrative synthesis was used to summarize the informing factors. Furthermore, Kirkpatrick's model was employed to categorize and summarize the effects of SA activities across various educational outcome levels, including reactions, learning, and behavioral changes.

Results: Our review reveals that key factors influencing students' SA includes feedback from various stakeholders, peer assessment, the format of SA, and both learner and teacher training. Among the 39 studies examined, the majority demonstrated positive effects of SA on aspects such as attitudes and skills. However, a minority found no significant associations, with these outcomes frequently linked to contextual variables and how SA was implemented.

Conclusions: Our review does not focus on SA accuracy or its role as an assessment method; instead it delves into the role of SA as an integral learning practice. We explored its influencing factors and its impact across various outcomes. The findings indicate that for SA to effectively enhance student learning outcomes, it must be implemented with ample support, clear guidelines, and within contexts that encourage feedback and reflection.

ARTICLE HISTORY
Received 13 May 2024
Revised 9 October 2024
Accepted 14 October 2024

KEYWORDS
Self-assessment; self-regulated learning; lifelong learning; feedback; medical education

Introduction

The importance of self-assessment (SA), especially for the purpose of professional self-regulation, is increasingly acknowledged in health professions education [1]. The literature on SA in health professions education has undergone an evolving process where it was initially considered as 'the involvement of learners in judging whether or not learner-identified standards have been met' [1,p.S46]. However, over the last decade, it has increasingly been challenged as an ill-defined construct that serves more purposes than merely comparing against certain standards, which has been often found inconsistent [2,3]. From the socio-cognitive perspective, SA is an integral element of self-regulated learning (SRL), and in recent years has come to be seen as an important lifelong-learning strategy [4]. According to Zimmerman [5], SRL is a dynamic process characterized by three phases: forethought, performance, and reflection. The final phase reflection involves students self-assessing their own performance, as well as making adjustments for future learning. Under that guidance, the conceptualization of SA has been consistently evolving. Eva and Regehr [2] defined it as 'a process of personal reflection…for the purposes of making judgements regarding one's

CONTACT Binbin Zheng binbin.zheng@usuhs.edu Building 53-105, 4301 Jones Bridge Road, Bethesda, MD 20814-4712, USA
Supplemental data for this article can be accessed online at https://doi.org/10.1080/07853890.2024.2421441.

own current level of knowledge, skills, and understanding…improve overall performance' (p.581). Similarly in general higher education, Yan and Brown [6] (2017) has also conceptualized SA as a process where learners first determine the relevant performance criteria, then seek feedback, self-reflect, and gradually calibrate their judgments. They emphasized the cyclical nature of this process insofar as this calibration can alter individuals' performance-criteria determinations, meaning that all the steps have to be repeated.

Literature has revealed the inconsistency between SA and actual performance metrics [1,7,8]. Students have the tendency to either overestimate or underestimate themselves due to personal bias, a lack of clearly established assessment parameters [9], students' unpreparedness to do so [10,11], or even students' psychological pressure to disguise the truth [12]. A critical review on SA[13] has called for less focus on looking into the consistency and summative form of SA, but more on its formative form and the associated cognitive and affective mechanisms. The same review study [13] argued the learning-oriented purpose of SA in that the ultimate goal of SA is not for the assessment's sake itself; but to inform adjustment which further lead to deepen learning and improve performance. It has been proposed that we should reframe SA as a core competency or skill [14], and as a process rather than an outcome for learning [15].

The effect of SA interventions on student learning has been examined by previous reviews in the field of general education over the past two decades e.g. [12,16,17]. While a generally positive effect on academic achievement was found across these meta-analyses, it was suggested that the effect varies across different contexts due to the complexity of the SA process. Yan et al. [17] identified a number of moderators influencing the effect of SA on learning outcomes, including educational levels, the complexity of the intervention, training received, among others. Their review called for appropriate and careful consideration of personal and contextual factors when designing and implementing SA interventions.

In health professions education, few systematic reviews of SA have been conducted in specialist fields such as surgery [18] and dentistry [19]. Nayar et al. [18] found that the accuracy of SA in general surgery and other surgical specialties is influenced by age, experience, and environment. In addition, the use of retrospective video playback is beneficial for learners' SA in surgical training. The other review [19] of SA in dental education suggested more attention should be given to teach students how to perform SA and called for more research looking into the impact of SA on outcomes. The only scoping review [20] conducted in medical education pointed out the misled emphasis on improving the accuracy of SA as the accuracy itself does not necessarily correlate with student learning; instead, research on SA should be guided towards SA *for* learning, which is, how learners achieve desired learning goals through the practice of understanding their current performance as compared to standards or rubrics for them to meet their goals. In fact, Eva et al. [21] conducted a longitudinal study which revealed that medical students' accuracy in SA did not improve, even after 2.5 years of experience and feedback in SA. While not specifically focusing on SA, one scoping review of self-monitoring [22] suggested that self-monitoring aligns with more objective performance measures and is beneficial for improving performance. They also distinguished self-monitoring from SA in the sense that self-monitoring is in-the-moment self-awareness with specific tasks and usually happens during the performance stage of Zimmerman's SRL phases.

Thus, as suggested by previous review and empirical research, our current systematic review will not focus on research looking just into the accuracy of SA, nor research using SA as an assessment method to evaluate the effectiveness of other programs. Instead, we will focus on reviewing existing literature on SA as a learning activity or learning skill development, or studies focused on SA *for* learning and SA *as* learning. In particular, this review aims to answer the following research questions:

1. What factors inform the implementation of medical students' SA practices?
2. How does SA practices contribute to various educational outcome levels?

Methods

A systematic review was performed based on the guidelines established by the PRIMSA criteria [23].

Search strategy and information sources

We employed Cook and West's [24,p.943] stepwise approach to conducting systematic reviews in medical education, including (1) defining a focused research question; (2) evaluating the appropriateness of a systematic review; (3) assembling a team and writing a protocol; (4) searching for eligible studies; (5) defining inclusion and exclusion criteria; (6) abstracting key information; and (7) analyzing and synthesizing results.

Our search was conducted in the following five databases: MEDLINE, APA PsycInfo, Web of Science, ERIC, and Embase. Our search terms were ((self-assessment OR 'self assessment' OR self-assess OR self-evaluation OR 'self evaluation' OR self-evaluate) AND ('medical education' OR 'medical students' OR 'clerkship' OR 'residents' OR 'interns')), with the specific search strings used in each database included in supplemental Appendix A. We restricted our search to peer-reviewed empirical journal articles, on the grounds that they had undergone a rigorous peer-review process. Articles in the five databases were searched from inception until October 2023. Our initial search generated a total of 12,274 studies; among them, 1761 were duplicates. After two researchers (QH and JL) independently screened the titles and abstracts of the remaining 10,513 studies and discussed any disagreements about whether they met the basic inclusion criteria mentioned above, an additional 10,401 were removed, leaving 112 articles for full-text review.

Eligibility criteria and data collection

During full-text review, three further inclusion criteria were adopted: (1) that the research have well-defined research questions as well as clear presentations of its methodology and findings, and (2) SA was the focus of the study, not just a datapoint for other outcomes. Each full text article was reviewed by two authors. Any disagreements were discussed as a group with all three authors (BZ, QH, and JL, all having a background in education and social science) until consensus was reached. At this stage, we excluded 70 articles that only used SA as a data-collection tool or only compared the accuracy of SA with other assessment methods, or described it in general terms, or were focused chiefly on curriculum design. This resulted in a final pool of 39 studies for analysis. Inter-rater reliability (Cohen's kappa) for the full-text review stage was calculated using Covidence software and was found to be 75%. The entire screening process was carried out using the Covidence systematic review management platform [25], which automatically created the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart that can be seen in Figure 1 [23].

Data analysis

First, an overall summary of the characteristics of each study – including country (countries) where the study was conducted, population, methodology and data collection, and themes based on the proposed research questions – was created. After that, qualitative narrative synthesis [10] was adopted to identify the major themes found across multiple studies, which involves the process of reading and re-reading the studies to identify similarities and differences, grouping findings into thematic categories, and integrating and narratively synthesizing themes. To answer the second research question regarding the effectiveness of SA activities, the Kirkpatrick model [26] was adopted, focusing on four levels of evaluation: Level 1 – reaction: assessment of learners' views; Level 2 – learning: changes in learners' knowledge and skills; Level 3 - – behavior: changes in learners' behaviors; Level 4 - – - outcome: changes in patient outcomes attributed to the learning intervention.

Findings

Study selection

A total of 39 studies were selected and included in our review.

Study characteristics

A detailed summary of the characteristics of the 39 reviewed articles is presented in the supplemental Appendix B. Among them, 19 studies were conducted in the United States, four studies were conducted in the United Kingdom, three each in Canada and India, and two studies were conducted in Australia. One study collected the data in two areas: the Netherlands and India. The seven countries and regions in which the remaining seven studies were conducted were Austria, Denmark, Taiwan, Malaysia, the Netherlands, Saudi Arabia, and the United Arab Emirates. In terms of the population, 27 of the studies focused on undergraduate medical education (UME), 8 focused on graduate medical education (GME), and an additional 4 focused on both UME and GME. Methodologically, 26 studies were quantitative, 7 qualitative, and 6 mixed-methods. For the themes of the included studies, we summarized four themes in total: (1) Informing factors of SA: Feedback and contextual influences; (2) Reaction-level outcomes: How students and faculty perceive SA; (3) Learning-level outcomes: Regular SA enhances knowledge and skill development; and (4) Behavior-level outcomes: Regular SA stimulates behavioral changes. In our review, we did not find any studies focusing on Level 4 which is the patient outcome, thus this theme was not included in our reporting of results. Furthermore, six studies collected data from multiple sites, while the other 33 collected data from a single

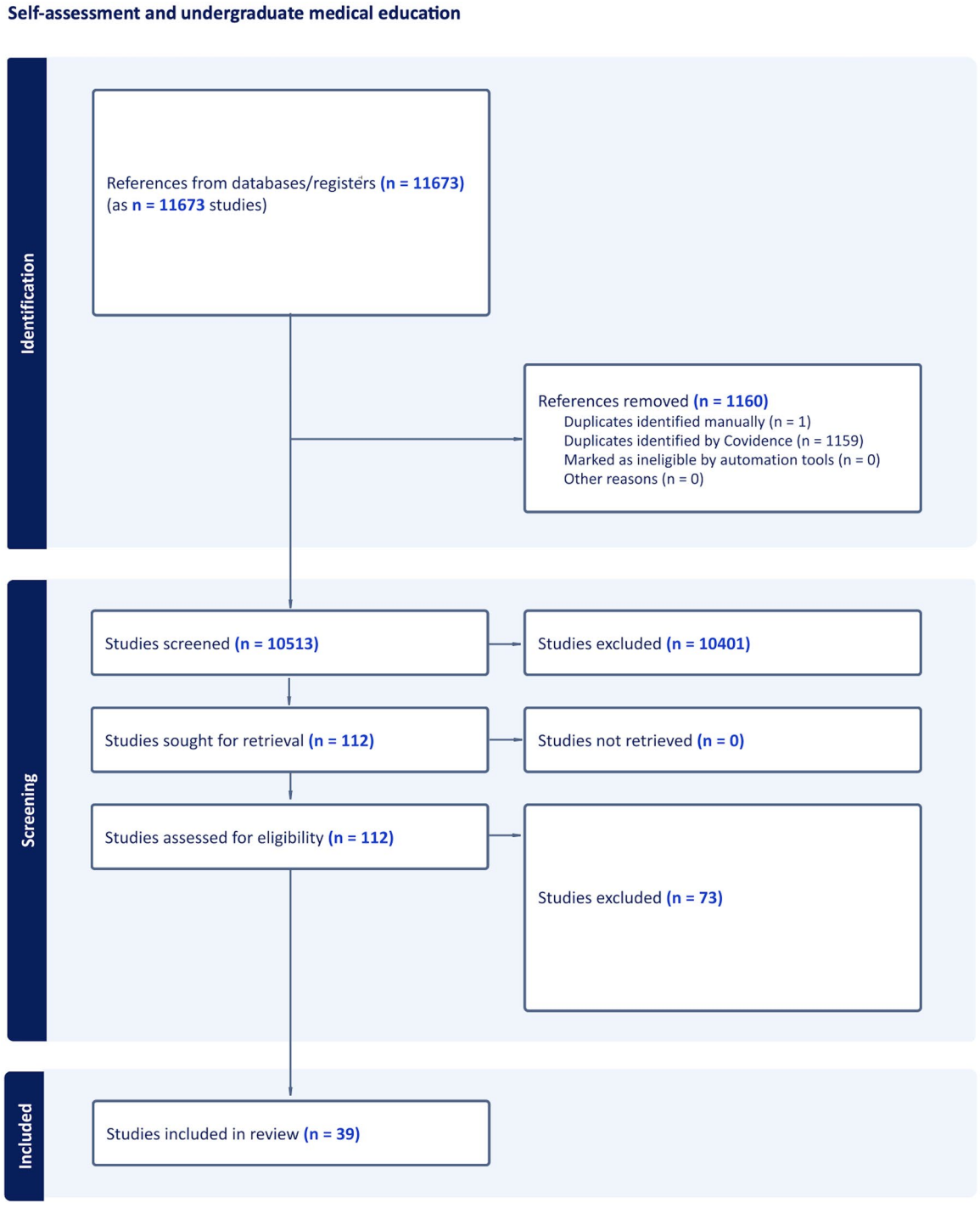

Figure 1. PRISMA flowchart for a systematic review of SA in undergraduate and graduate health-professions education.

site. To assess study quality, we adapted the Critical Appraisal Skills Programme (CASP) checklist [27], which consists of 10 items evaluating various aspects of a study, including the appropriateness of the methodology, research design, and recruitment strategy, the adequacy of data analysis, and so on. Studies that received eight or more 'yes' responses were deemed strong, while those with six to seven 'yes' answers were considered moderate. In our analysis, 34 studies were classified as strong, while the remaining 5 were classified as moderate (see supplemental Appendix B).

Informing factors of SA: feedback and contextual influences

We have summarized four themes of informing factors of SA, see Figure 2 for an overview.

Multi-source feedback

In terms of factors informing the implementation of medical students' SA activities, the sampled literature frequently recommended that existing benchmarks [28] or feedback – whether from existing benchmarks, teachers [29–33], peers [34], or automated online SA tools [35] – should be coupled with learners' SA. For example, Hawkins and colleagues [28] found that video-recorded feedback together with standardized video-recorded benchmarks could improve students' SA accuracy of their clinical skills. The same authors stressed the importance of providing explicit standards because, unless this was done, inexperienced learners would instead use implied benchmarks they had developed themselves. In their 2023 study, Robb and Rosenbaum [36] examined the interplay between SA and feedback through the perspective of situated learning theory. They found that when feedback givers ask learners to engage in SA before providing feedback, it facilitates a more productive dialogue between learners and feedback providers. However, they noted a caveat: unprepared learners might approach SA with a focus on social desirability, prioritizing what they believe the feedback provider wants to hear, rather than an honest reflection of their own performance. Furthermore, in terms of the quality of feedback, one study that looked into the relationship between SA and the nature of feedback found that high-quality positive feedback was associated with students' overly optimistic SAs of their performance, whereas high-quality negative feedback was linked to more accurate SA [37]. Similarly, Elnicki and Zalenski [38] concluded that students struggled with SA regardless of whether it was self-initiated or performed in response to teachers' feedback, and highlighted the importance of well-articulated, specific feedback by preceptors.

SA and peer assessment

Some other studies in our review recommended that SA be paired with peer assessment to improve students' self-directed learning and performance. Emke et al. [39] mentioned that, in the case of certain

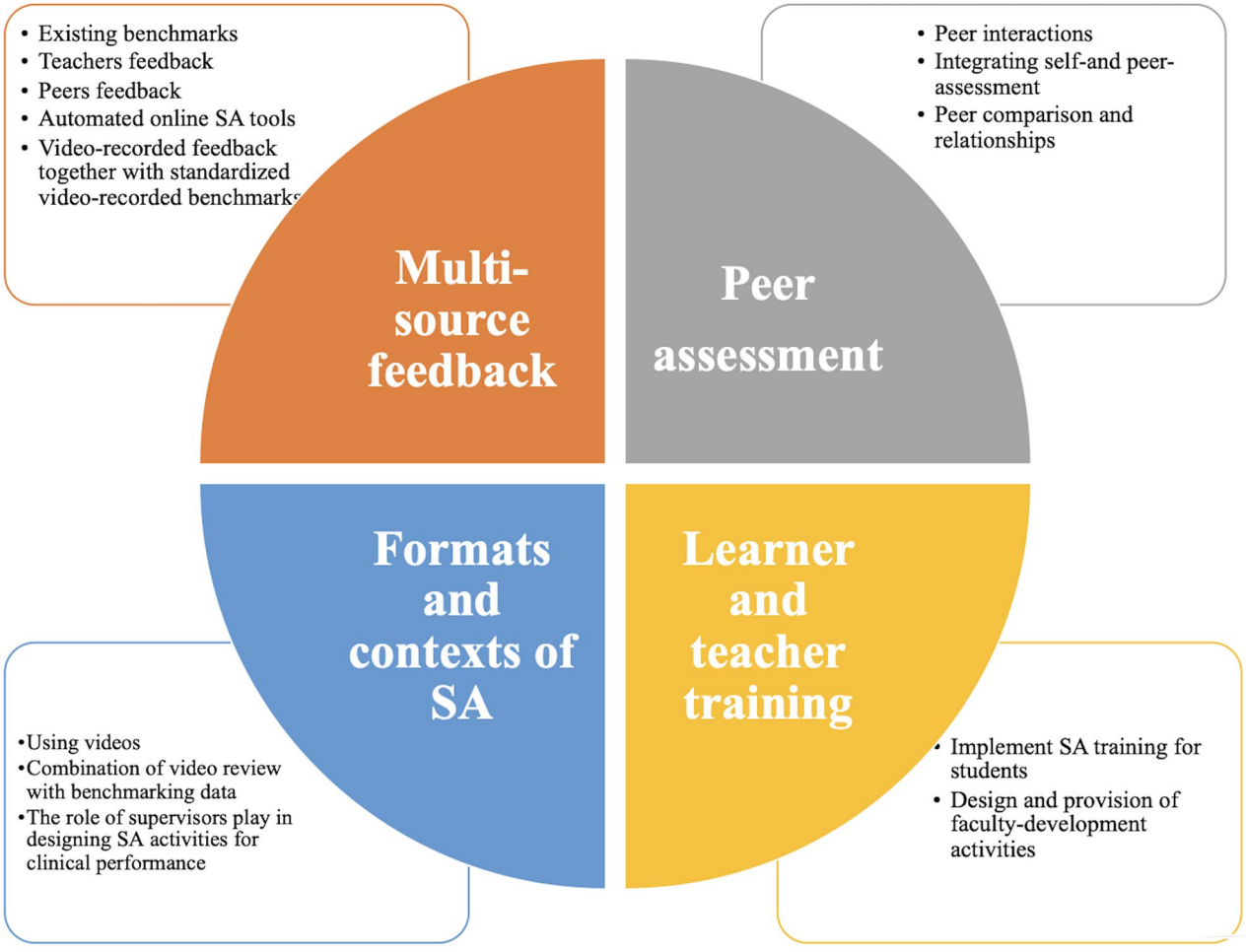

Figure 2. Informing factors of self-assessment. The figure illustrates four informing factors of self-assessment along with relevant related factors.

behavioral competencies that may not be directly observed by supervisors (e.g. interpersonal skills, communication, teamwork), peer assessment might be the better of these two mechanisms, on the grounds that peers have more direct interactions with each other. The same study, in line with other literature [40], suggested that SA should not be used in isolation, but rather should be bundled with other activities to promote multi-source feedback. Bryan et al. [41] also found support for the combined use of peer and SA in enhancing medical students' assessment abilities. Notably, these students tended to be more self-critical compared to their evaluation of peers. The study emphasized that the primary goal of integrating self- and peer-assessment is to train students in effectively evaluating their own and others' professional behaviors. This process involves reflective judgment rather than real-time observation and tracking of behaviors, and it intended to cultivate a deeper understanding of professional conduct rather than serve as an exact measure of professionalism. Even without direct peer assessment, the aspects of peer comparison and relationships contribute to supporting informed SA. This is achieved by fostering a sense of belonging among peers [42].

Formats and contexts of SA

Regarding the formats of SA activities, several of the reviewed papers suggested that using videos could be an effective means of helping students make more accurate SAs, especially of their clinical skills [9,43]. Contrastingly, another two studies [44,45] suggested that video review alone might not enhance learners' SA. Srinivasan et al. [44] further argued that it is the combination of video review with benchmarking data that could be effective, likely by aiding learners in interpreting their performance against objective standards.

Sargeant et al. [32], through a qualitative study with 134 participants across the spectrum of medical education, conceptualized SA as a multifaceted construct shaped by various internal factors (e.g. emotion, experience, confidence), process factors (e.g. instructional activities, guidelines, assessments), and people (e.g. supervisors, peers, patients). In their other study [46], they further stressed the important role supervisors play in designing SA activities for clinical performance that they need to be: (1) *prepared*, especially in the forms of setting specific standards and learning feedback-provision skills; and (2) *engaged*, notably in the forms of taking an interest in learners and sharing performance information with them regularly. However, the same study pointed out that whether these factors effectively inform students' SA or not is dependent on personal and contextual conditions. In other words, it is not so much the approach itself as how it is used that influences the effectiveness of activities for informing SA. Nevertheless, the authors conceded the importance of regular verbal feedback and of striking a balance between teacher guidance/feedback and students' self-directed learning.

Learner and teacher training

In addition to those contextual factors mentioned above for informing learners' SA practices, the included studies also argued that explicit training aimed at enhancing students' SA skills ought to be more widely provided [40]. Caverzagie et al. [47] observed that, due to insufficient training in SA, residents primarily identified broad and unspecific learning objectives in areas like medical knowledge and patient care, while neglecting other competencies. Furthermore, Rees & Shepherd [48] suggest that while SA training should address its educational benefits and challenges, offering learners continuous opportunities to practice these skills and negotiate grades with peers and instructors could further develop their lifelong SA abilities. Additionally, two studies called for the design and provision of faculty-development activities not just provided for learners, but instead, aimed at raising instructors' awareness of the importance of learners' SA, as well as how to scaffold it [46,49].

Reaction-level outcomes: how students and faculty perceive SA

According to the studies that examined perceptions towards SA, both students and faculty members value the experience of SA regarding attitudinal and behavioral changes in students [50,51]. Particularly, students believe the process of SA is helpful to increasing their knowledge and interest in learning which further motivate them to develop self-directed learning skills [50]. Students also think SA leads to gaining long-term knowledge and active engagement in learning activities [51], improved self-awareness of their learning needs [52], increased confidence and less anxiety in dealing with patients and professionals, and higher self-acceptance [53]. However, Al-Kadri et al. [40] noted that students did *not* perceive SA as a driver of their learning, probably because they perceive SA practices as isolated from other components, such as supervisors' feedback, of their education. The same study also reported that, while students held favorable views of the training in SA, they did not perceive such training

as strategically helpful to their practice or even to their SA accuracy. Interestingly, students admitted that they believe overestimation in SA can manipulate supervisors to give higher marks. While few studies looked into faculty's reactions to learners' SA, one study mentioned that faculty members also perceive SA as an effective tool for improving students' self-directed learning, but they are recommended to play a more active role in motivating students to self-assess [50].

Despite many positive perceptions presented, some negative perceptions to SA were also reported, mostly by students: e.g. that it is time-consuming [50], arouses uncomfortable feelings when based on peers' feedback, and lack of privacy when conducting SA around peers [54]. (See Figure 3 for an overall mapping of impact of SA on different levels of educational outcomes)

Learning-level outcomes: regular SA enhances knowledge and skill development

The studies focusing on examining the effect of SA on learning could be categorized into the following four outcomes: knowledge acquisition, clinical performance, non-technical skills, and overall learning skills.

Winton et al. [55] detected a positive association between participation in weekly SA and knowledge retention from grand round presentations among medical students and residents. Similarly, Sharma et al. [50] reported that SA helped first-year medical students improve their examination scores. On the contrary, there were also two studies [38,56] suggesting a scant association between SA and academic performance. Elnicki and Zalenski further explained that SA is a difficult exercise for students, thus proposed a need to structure SA to help learners to specify their strengths and weaknesses in clinical, knowledge, and attitudinal domains.

SA was also recognized for enhancing students' clinical performance, particularly in the performance of surgical training [43,57–59]. For example, Andersen et al. [57] revealed that students who engaged in structured SA demonstrated improved performance in virtual reality surgical procedures and achieved higher passing rates compared to their counterparts. Similarly, Palvia et al. [43] concluded that video SA can be effectively incorporated into surgical training programs and learners who participated in video SA improved more significantly in ped and needle drops between sessions. Only one study focused on a specific training course in SA in laparoscopic surgical skills [59]. Their study revealed the effectiveness of SA training for improving surgical residents' overall surgical performance, and emphasized the importance of 'reflection-*before*-practice' in the development of surgical skills. However, one study found no relationship. Lin et al. [60] suggested that core self-evaluation tendencies positively influenced workplace compassion satisfaction and mitigated burnout, although no

Figure 3. Impact of self-assessment on educational outcome. This figure describes the impact of self-assessment (positive, neutral, and negative) on three educational outcomes, including reaction, learning, and behavior levels.

significant effects were observed on clinical competence during the clerkship period.

After engaging in SA activities, medical students demonstrated better interviewing skills [31,61] and communication skills (e.g. building rapport with peers, active listening) [54]. Hanley et al.'s study [31], in which SA was implemented together with goal setting, indicated that first-year medical students who had good SA and goal-setting skills demonstrated an increase in interviewing skills. In Lane and Gottlieb's study [61], a self-study and SA manual was developed to facilitate students' reviewing of their own patient encounter videos. The result found improved interpersonal skills, SA skills, and student performance of history taking. In addition, by comparing self-assessed and external procedure-based assessment on changes in clinical practice, another study [62] suggested that while technical skills were not different between these two groups, use of SA was more closely related to situation awareness, decision-making, and leadership.

Finally, three studies [63–65] examined how SA contributed to students' development of learning goals, which is a SRL skill related to goal setting. All of them suggested that learners created their learning goals based on SA, and they were more likely to act on their learning goals when feedback is incorporated. Wolff et al. [65] further concluded in their study on pre-clinical medical students that those who had higher SA were more likely to act on their learning goals, which emphasized the importance of SA not only on goal setting, but also on learners' subsequent actions.

Behavior-level outcomes: regular self-assessment stimulates behavioral changes

We also found a total of four studies looking into the correlations between SA and behavioral change. For example, Gruppen et al. [66] found that although medical students' SA of their diagnostic skills did not correlate with their allocation of study time, which is an indication for self-directed learning, there is are either positive or negative relationships between time allocation and students' self-assessed strength and weakness, in the way that some students tended to spend more time on areas they considered weak, while some tended to spend less time on self-assessed weak areas.

Andersen et al. [57] conducted a comparison study and found that, medical students who conducted SA after each simulation-based surgical procedure had significantly less number of repetitions needed than their counterparts who did not conduct SA. Similarly, a randomized control study found that the video SA group had both shorter completion time as well as improved pre- to post-completion time than the control group who did not engage in video SA in two laparoscopic simulation training tasks [43].

Only one study examined SA and mental health outcomes such as wellbeing. Lin et al. [60] suggested that core self-evaluation, defined as a general skill about a person's evaluation of their worth, competence, and capability, had significantly positive effects on medical students' wellbeing, as measured by their compassion satisfaction and reversed burnout.

Discussion

The review conducted systematically explored research on SA within both undergraduate and graduate medical education contexts. This section delves into the outcomes of that exploration, discussing the implications and findings from dual viewpoints: firstly, SA's role as a pedagogical strategy for enhancing learning outcomes, and secondly, its significance in fostering skill development for lifelong learning.

SA as a pedagogical strategy: enhancing learning

Yan and Brown [6] have articulated that the primary aim of SA transcends merely training students to accurately assess themselves; its core intention is to serve formative functions, specifically to bolster learning. Our review indicates that SA, on its own, might not significantly impact students' academic performance [38,56]. For SA to effectively contribute to improving learning outcomes, it should be integrated as part of a broader formative educational framework, including combining SA with reflective practice, goal-setting exercises, peer feedback, feedback from other stakeholders such as preceptors or simulated patients [33], and potentially the utilization of video recordings [31,54]. Such a multifaceted approach aligns with findings from other domains, demonstrating that SA, when guided by explicit criteria, paired with strategic educational interventions, and complemented by feedback and reflective practices, can enhance not only students' clinical performance but also their SA skills [67]. Moreover, in instances where studies did not find a direct correlation between SA engagements and learning enhancements, it was often pointed out that additional support in structuring SA processes could further assist students. This implies a need for more comprehensive scaffolding around SA practices, ensuring students are well-equipped to engage in these activities effectively [38].

The significance of feedback in medical education is a well-documented phenomenon, underscoring its pivotal role in shaping educational outcomes [68,69], Similarly, SA garners attention within educational studies, particularly due to its integral position within the self-regulated learning framework. Our review underscores the critical role of feedback within SA practices, illuminating how it enriches the SA experience. Robb et al. [36] proposed that incorporating SA into the feedback conversation could not only engage learners but also empower them in both giving and receiving feedback; however, a safe environment should be created to ensure learners express their true perceptions of their performance.

Teacher feedback serves as a crucial benchmark, enabling students to align their self-assessments with established standards, thus enhancing the validity of SA [28]. Although peer feedback and peer assessment have the merits of more direct interaction than teacher-student interaction can generally provide, these activities need to be improved in terms of both their specificity and how actionable they are [34]. While SA and peer-assessment may not be reliable or valid as the basis of summative assessment, they stand out as instrumental strategies for improving student learning [70] and to promote metacognitive awareness and autonomous learning, which are of vital importance to health professions students' future professional endeavors [71]. A meta-analysis in higher education [17] reveals that although SA and peer-assessment have overall positive effect sizes on academic performance, the implementation of SA and peer-assessment activities are often complex and influenced by a wide range of personal and contextual factors such as students' readiness or the support provided.

This discussion highlights that engaging in SA enables a deeper comprehension of learning objectives, criteria, and methodologies. This enhanced understanding fosters confidence, experience accumulation, and insightful reflections on the educational journey [72,73]. Moreover, by facilitating self-identification, SA can propel students towards self-directed learning. Hence, recognizing SA as a foundational skill within health professions education and ensuring its explicit instruction is paramount for preparing students for their forthcoming professional roles [50].

Pursuit of accuracy in SA for calibration and future improvement

It is important to acknowledge that while SA has been criticized for its potential misalignment with other assessment methods – and therefore should not be solely relied upon for any high-stakes decision-making [13,74] – our study intentionally does not focus on the accuracy of SA for such decisions. However, our review underscores the inherent value in the pursuit of accuracy within the context of SRL.

According to Yan and Brown's [6] SA model, students use SA in conjunction with external feedback to calibrate their performance. This calibration process is crucial as it enables students to more accurately identify their strengths and weaknesses, thereby facilitating targeted improvements in learning. Moreover, engaging in this cyclical process of calibration, where learners continuously refine their SA by determining criteria, receiving feedback, and adjusting their self-assessment, fosters a deeper understanding of their learning process.

Thus, while the alignment of SA with other assessment methods is not the primary focus and should not be the goal, the process of seeking accuracy in SA remains a valuable exercise for enhancing students' learning outcomes and fostering a more reflective and informed approach to self-regulation.

SA as a mechanism for enhancing Self-regulated learning

Engaging students in SA aims not only to elevate their academic performance but also to foster academic self-regulation, or the capacity to oversee their own learning trajectory [75,76]. SA stands as a crucial component of the self-regulated learning process, underpinning various SRL behaviors [77,78]. That is, it plays a predictive role during the pre-learning stage; a concurrent one during the performance stage; and a summative role at the end of learning [1]. Research in higher education underscores that SA is a continuous element throughout the SRL journey, positioning it more as a strategic approach to learning rather than a mere evaluative tool. The incorporation of SA at multiple points within the SRL cycle underscores the necessity for students to develop SA competencies early in their learning endeavors, rather than deferring this skill acquisition until the end of the learning process [79].

Nonetheless, our analysis indicates a scarcity of studies explicitly addressing the interplay between SA and SRL as a facet of skill development [48,65]. Among the limited research, Rees & Shepherd [48] advocate for recognizing the dual value of SA: its educational merits and its potential in nurturing lifelong learning capabilities. The ability to discern personal strengths and limitations is instrumental for students in formulating learning objectives and fostering enduring learning habits [6]. However, the development of SA skills should not be left to chance but rather should be supported through deliberate instructional strategies [40].

Our review also suggests that SA skills can be cultivated by prompting students to reflect on their video-recorded clinical practice [28,31], to compile their own learning portfolios [80], to engage in peer-assessment and peer-feedback exercises [28,39], and/or to enroll in courses specifically designed to learn how to learn [81].

Implications for future research

Although the correlation between SA and other measures of academic performance has often been characterized as weak, SA remains a vital element in medical education, enabling students to acknowledge and articulate their learning gaps [21]. Future research could pivot from solely assessing the accuracy of SA in comparison to other evaluative measures. Given the personalized and contextual nature of SA, investigations could broaden to explore how individual, environmental, and cultural dynamics influence SA practices [36,40,46]. Additionally, research could aim to identify effective strategies for cultivating students' SA competencies, thus equipping them with the skills necessary for continuous learning and professional growth. Finally, we observed that the majority of the included studies were conducted in the Global North. Conducting more empirical studies in the Global South would provide valuable insights and help gain a more diverse perspective on this topic.

Limitations

Our study has several limitations. First, our search was conducted using keywords related to SA and self-evaluation, but did not include terms such as self-reflection, self-judgment, or self-monitoring, which some authors may use interchangeably with SA. Consequently, this may have led to the exclusion of certain relevant articles. However, the initial search generated a substantial number of references, which may help mitigate this limitation. Second, while SA is often used alongside peer assessment or 360-degree assessment to examine the consistency of different assessments, our choice to exclude articles focused solely on SA accuracy may limit the comprehensiveness of our review. The decision was based on the view that SA is more of a practice and skill than purely an assessment method.

Conclusions

SA is arguably essential to learners' learning and their careers beyond the educational training [82]. Understanding their own strengths and weaknesses can help learners focus on what to improve and direct their learning into more effective directions, fostering SRL abilities that are important life-long learning skills [83]. This systematic review, exploring the implementation of SA as a pedagogical approach and skill development, advocates for viewing SA not merely as an assessment method but as a key learning activity and a critical skill for supporting ongoing professional growth and lifelong learning.

Author contributions

Binbin Zheng contributed to the conception of the study, literature search, analysis and interpretation of data, drafting and revising of the manuscript, and the final approval of the version to be published.

Qing He contributed to the literature search, analysis and interpretation of data, drafting and revising of the manuscript, and the final approval of the version to be published.

Junru Lei contributed to the literature search, analysis and interpretation of data, drafting of the manuscript, and the final approval of the version to be published.

All authors agree to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Disclaimer

The opinions and assertions expressed herein are those of the authors and do not reflect the official policy or position of the Uniformed Services University of the Health Sciences or the Department of Defense.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Binbin Zheng http://orcid.org/0000-0001-6160-0104
Qing He http://orcid.org/0000-0003-4260-7997
Junru Lei http://orcid.org/0000-0002-1841-7758

Data availability statement

Data will be shared upon reasonable request made to the corresponding author of the manuscript.

References

[1] Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med.

2005;80(10 Suppl):S46–S54. doi: 10.1097/00001888-200510001-00015.
[2] Eva KW, Regehr G. Knowing when to look it up: a new conception of self-assessment ability. Acad Med. 2007; 82(10 Suppl):S81–S84. doi: 10.1097/ACM.0b013e31813e6755.
[3] Eva KW, Regehr G, Gruppen LD. Blinded by "insight": self-assessment and its role in performance improvement. In: The question of competence. Ithaca (NY): Cornell University Press; 2012. p. 131–54. Available from: https://www.degruyter.com/document/doi/10.7591/9780801465802-010/pdf?licenseType=restricted.
[4] Zimmerman BJ. Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. Am Educ Res J. 2008;45(1):166–183. doi: 10.3102/0002831207312909.
[5] Zimmerman BJ. Becoming a self-regulated learner: an overview. Theory Pract. 2002;41(2):64–70. doi: 10.1207/s15430421tip4102_2.
[6] Yan Z, Brown GTL. A cyclical self-assessment process: towards a model of how students engage in self-assessment. Assess Eval High Educ. 2017;42(8):1247–1262. doi: 10.1080/02602938.2016.1260091.
[7] Brown GTL, Andrade HL, Chen F. Accuracy in student self-assessment: directions and cautions for research. Assess Educ Princ Policy Pract. 2015;22(4):444–457. doi: 10.1080/0969594X.2014.996523.
[8] Baxter P, Norman G. Self-assessment or self deception? A lack of association between nursing students' self-assessment and performance. J Adv Nurs. 2011;67(11): 2406–2413. doi: 10.1111/j.1365-2648.2011.05658.x.
[9] Sanderson TR, Kearney RC, Kissell D, et al. Evaluating student self-assessment through video-recorded patient simulations. J Dent Hyg. 2016;90(4):257–262.
[10] Mould MR, Bray KK, Gadbury-Amyot CC. Student self-assessment in dental hygiene education: a cornerstone of critical thinking and problem-solving. J Dent Educ. 2011;75(8):1061–1072.
[11] Shaban S, Aburawi E, Elzubeir K, et al. Factors influencing medical students' self-assessment of examination performance accuracy: a United Arab Emirates study. Educ Health (Abingdon). 2016;29(2):75–81. doi: 10.4103/1357-6283.188688.
[12] Harris LR, Brown GTL. Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: case studies into teachers' implementation. Teach Teach Educ. 2013;36:101–111. doi: 10.1016/j.tate.2013.07.008.
[13] Andrade HL. A critical review of research on student self-assessment. Front Educ. 2019;4:87. Available from: https://www.frontiersin.org/articles/10.3389/feduc.2019.00087.
[14] Brown GTL, Harris LR. The future of self-assessment in classroom practice: reframing self-assessment as a core competency. Frontline Learn Res. 2014;2(1):22–30.
[15] Yan Z, Panadero E, Wang X, et al. A systematic review on students' perceptions of self-assessment: usefulness and factors influencing implementation. Educ Psychol Rev. 2023;35(3):81. doi: 10.1007/s10648-023-09799-1.
[16] Topping K. Self and peer assessment in school and university: reliability, validity and utility. In: Segers M, Dochy F, Cascallar E, editors. Optimising new modes of assessment: in search of qualities and standards. Dordrecht: Springer Netherlands; 2003. p. 55–87. doi: 10.1007/0-306-48125-1_4.
[17] Yan Z, Lao H, Panadero E, et al. Effects of self-assessment and peer-assessment interventions on academic performance: a meta-analysis. Educ Res Rev. 2022;37:100484. doi: 10.1016/j.edurev.2022.100484.
[18] Nayar SK, Musto L, Baruah G, et al. Self-assessment of surgical skills: a systematic review. J Surg Educ. 2020;77(2):348–361. doi: 10.1016/j.jsurg.2019.09.016.
[19] Mays KA, Branch-Mays GL. A systematic review of the use of self-assessment in preclinical and clinical dental education. J Dent Educ. 2016;80(8):902–913.
[20] Yates N, Gough S, Brazil V. Self-assessment: with all its limitations, why are we still measuring and teaching it? Lessons from a scoping review. Med Teach. 2022;44 (11):1296–1302. doi: 10.1080/0142159X.2022.2093704.
[21] Eva KW, Cunnington JPW, Reiter HI, et al. How can I know what I don't know? Poor self assessment in a well-defined domain. Adv Health Sci Educ Theory Pract. 2004;9(3):211–224. doi: 10.1023/B:AHSE.0000038209.65714.d4.
[22] Johnson WR, Durning SJ, Allard RJ, et al. A scoping review of self-monitoring in graduate medical education. Med Educ. 2023;57(9):795–806. doi: 10.1111/medu.15023.
[23] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: 10.1136/bmj.n71.
[24] Cook DA, West CP. Conducting systematic reviews in medical education: a stepwise approach. Med Educ. 2012; 46(10):943–952. doi: 10.1111/j.1365-2923.2012.04328.x.
[25] Covidence systematic review software [Internet]. Melbourne, Australia: Veritas Health Innovation. Available from: www.covidence.org.
[26] Kirkpatrick D, Kirkpatrick J. Evaluating training programs: the four levels. Okaland (CA): Berrett-Koehler Publishers; 2006. p. 399.
[27] CASP Checklists - Critical Appraisal Skills Programme. [Internet]. CASP - Critical Appraisal Skills Programme. [cited 2024 Aug 30]. Available from: https://casp-uk.net/casp-tools-checklists/.
[28] 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):279–284. doi: 10.3109/0142159X.2012.658897.
[29] Lundquist LM, Shogbon AO, Momary KM, et al. A comparison of students' self-assessments with faculty evaluations of their communication skills. Am J Pharm Educ. 2013;77(4):72–72. doi: 10.5688/ajpe77472.
[30] Torres MBMD, Cochran AMD. Accuracy and content of medical student midclerkship self-evaluations. Am J Surg. 2016;211(6):1153–1157. doi: 10.1016/j.amjsurg.2015.11.030.
[31] Hanley K, Zabar S, Charap J, et al. Self-assessment and goal-setting is associated with an improvement in interviewing skills. Med Educ Online. 2014;19(1):24407–24407. doi: 10.3402/meo.v19.24407.
[32] Sargeant J, Armson H, Chesluk B, et al. The processes and dimensions of informed self-assessment: a conceptual model. Acad Med. 2010;85(7):1212–1220. doi: 10.1097/ACM.0b013e3181d85a4e.
[33] Tanoue N, Korovin LN, Carton M, et al. Faculty feedback versus residents' self-assessment of operative performance:

different but complementary. Am J Surg. 2018;215(2):288–292. doi: 10.1016/j.amjsurg.2017.11.016.
[34] Hulsman RL, van der Vloodt J. Self-evaluation and peer-feedback of medical students' communication skills using a web-based video annotation system. Exploring content and specificity. Patient Educ Couns. 2015;98(3):356–363. doi: 10.1016/j.pec.2014.11.007.
[35] Jagadeesan AR, Neelakanta RR. Online self-assessment tool in biochemistry–a medical student's perception during COVID-19 pandemic. J Educ Health Promot. 2021;10:1–6.
[36] Robb KA, Rosenbaum ME, Peters L, et al. Self-assessment in feedback conversations: a complicated balance. Acad Med. 2023;98(2):248–254. doi: 10.1097/ACM.0000000000004917.
[37] Plakht Y, Shiyovich A, Nusbaum L, et al. The association of positive and negative feedback with clinical performance, self-evaluation and practice contribution of nursing students. Nurse Educ Today. 2013;33(10):1264–1268. doi: 10.1016/j.nedt.2012.07.017.
[38] Elnicki DM, Zalenski D. Integrating medical students' goals, self-assessment and preceptor feedback in an ambulatory clerkship. Teach Learn Med. 2013;25(4):285–291. doi: 10.1080/10401334.2013.827971.
[39] Emke AR, Cheng S, Chen L, et al. A novel approach to assessing professionalism in preclinical medical students using multisource feedback through paired self- and peer evaluations. Teach Learn Med. 2017;29(4):402–410. doi: 10.1080/10401334.2017.1306446.
[40] Al-Kadri H, Al-Moamary MS, Al-Takroni H, et al. Self-assessment and students' study strategies in a community of clinical practice: a qualitative study. Med Educ Online. 2012;17(1):11204–11204. doi: 10.3402/meo.v17i0.11204.
[41] Bryan RE, Krych AJ, Carmichael SW, et al. Assessing professionalism in early medical education: experience with peer evaluation and self-evaluation in the gross anatomy course. Ann Acad Med Singap. 2005;34(8):486–491. doi: 10.47102/annals-acadmedsg.V34N8p486.
[42] McDonald J, Ryan S, Heeneman S, et al. Informed self-assessment during the transition to medical school: a longitudinal narrative study. BMJ Open. 2022;12(12):e065203. doi: 10.1136/bmjopen-2022-065203.
[43] Palvia V, Huntly J, Seckin S, et al. Role of video self-assessment in laparoscopic simulation training: a randomized pilot trial. AJOG Glob Rep. 2023;3(3):100224. doi: 10.1016/j.xagr.2023.100224.
[44] Srinivasan M, Hauer KE, Der-Martirosian C, et al. Does feedback matter? Practice-based learning for medical students after a multi-institutional clinical performance examination. Med Educ. 2007;41(9):857–865. doi: 10.1111/j.1365-2923.2007.02818.x.
[45] Vnuk A, Owen H, Plummer J. Assessing proficiency in adult basic life support: student and expert assessment and the impact of video recording. Med Teach. 2006;28(5):429–434. doi: 10.1080/01421590600625205.
[46] 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):636–647. doi: 10.1111/j.1365-2923.2010.03888.x.
[47] Caverzagie KJ, Shea JA, Kogan JR. Resident identification of learning objectives after performing self-assessment based upon the ACGME core competencies. J Gen Intern Med. 2008;23(7):1024–1027. doi: 10.1007/s11606-008-0571-7.
[48] Rees C, Shepherd M. Students' and assessors' attitudes towards students' self-assessment of their personal and professional behaviours. Med Educ. 2005;39(1):30–39. doi: 10.1111/j.1365-2929.2004.02030.x.
[49] Elhadi M, Ahmed H, Khaled A, et al. Informed self-assessment versus preceptor evaluation: a comparative study of pediatric procedural skills acquisition of fifth year medical students. BMC Med Educ. 2020;20(1):318–318. doi: 10.1186/s12909-020-02221-2.
[50] Sharma R, Jain A, Gupta N, et al. Impact of self-assessment by students on their learning. Int J Appl Basic Med Res. 2016;6(3):226–229. doi: 10.4103/2229-516X.186961.
[51] Melser CM, Lettner S, Bäwert A, et al. Pursue today and assess tomorrow - how students' subjective perceptions influence their preference for self- and peer assessments. BMC Med Educ. 2020;20(1):479–479. doi: 10.1186/s12909-020-02383-z.
[52] Kennedy G, Rea JNM, Rea IM. Prompting medical students to self-assess their learning needs during the ageing and health module: a mixed methods study. Med Educ Online. 2019;24(1):1579558–1579558. doi: 10.1080/10872981.2019.1579558.
[53] Kennell JH, Tempio CR, Wile MZ. Self-evaluation by first-year medical students in a clinical science programme. Br J Med Educ. 1973;7(4):230–238. doi: 10.1111/j.1365-2923.1973.tb02239.x.
[54] Perera J, Mohamadou G, Kaur S. The use of objective structured self-assessment and peer-feedback (OSSP) for learning communication skills: evaluation using a controlled trial. Adv Health Sci Educ Theory Pract. 2010;15(2):185–193. doi: 10.1007/s10459-009-9191-1.
[55] Winton LM, Ferguson EMN, Hsu C-H, et al. Does self-assessment improve the effectiveness of grand rounds lectures in a community-based teaching hospital? J Surg Educ. 2016;73(6):968–973. doi: 10.1016/j.jsurg.2016.04.014.
[56] Yoho RM, Vardaxis V, Millonig K. Relationship between academic performance and student self-assessment of clinical performance in the College of Podiatric Medicine and Surgery. J Am Podiatr Med Assoc. 2016;106(3):214–217. doi: 10.7547/14-166.
[57] Andersen SAW, Guldager M, Mikkelsen PT, et al. The effect of structured self-assessment in virtual reality simulation training of mastoidectomy. Eur Arch Otorhinolaryngol. 2019;276(12):3345–3352. doi: 10.1007/s00405-019-05648-6.
[58] Karnick A, Limberg J, Bagautdinov I, et al. Can general surgery interns accurately measure their own technical skills? Analysis of cognitive bias in surgical residents' self-assessments. Surgery. 2021;170(5):1353–1358. doi: 10.1016/j.surg.2021.04.008.
[59] Ganni S, Botden SMBI, Schaap DP, et al. "Reflection-before-practice" improves self-assessment and end-performance in laparoscopic surgical skills training. J Surg Educ. 2018;75(2):527–533. doi: 10.1016/j.jsurg.2017.07.030.
[60] Lin YK, Chen DY, Lin BYJ. Determinants and effects of medical students' core self-evaluation tendencies on clinical competence and workplace well-being in clerkship.

PLoS One. 2017;12(11):e0188651. doi: 10.1371/journal.pone.0188651.
[61] Lane JL, Gottlieb RP. Improving the interviewing and self-assessment skills of medical students: is it time to readopt videotaping as an educational tool? Ambul Pediatr. 2004;4(3):244–248. doi: 10.1367/A03-122R1.1.
[62] Osborne AJ, Hawkins SC, Pournaras DJ, et al. An evaluation of operative self-assessment by UK postgraduate trainees. Med Teach. 2014;36(1):32–37. doi: 10.3109/0142159X.2013.836268.
[63] Bounds R, Bush C, Aghera A, et al. Emergency medicine residents' self-assessments play a critical role when receiving feedback. Acad Emerg Med. 2013;20(10):1055–1061. doi: 10.1111/acem.12231.
[64] Eva KW, Munoz J, Hanson MD, et al. Which factors, personal or external, most influence students' generation of learning goals? Acad Med. 2010;85(10 Suppl):S102–S105. doi: 10.1097/ACM.0b013e3181ed42f2.
[65] Wolff M, Stojan J, Cranford J, et al. The impact of informed self-assessment on the development of medical students' learning goals. Med Teach. 2018;40(3):296–301. doi: 10.1080/0142159X.2017.1406661.
[66] Gruppen LD, White C, Fitzgerald JT, et al. Medical students' self-assessments and their allocations of learning time. Acad Med. 2000;75(4):374–379. doi: 10.1097/00001888-200004000-00018.
[67] Poirier TI, Pailden J, Jhala R, et al. Student self-assessment and faculty assessment of performance in an interprofessional error disclosure simulation training program. Am J Pharm Educ. 2017;81(3):54–54. doi: 10.5688/ajpe81354.
[68] Bing-You RG, Trowbridge RL. Why medical educators may be failing at feedback. JAMA J Am Med Assoc. 2009;302(12):1330–1331. doi: 10.1001/jama.2009.1393.
[69] Ende J. Feedback in clinical medical education. JAMA. 1983;250(6):777–781. doi: 10.1001/jama.1983.03340060055026.
[70] Roberts C, Jorm C, Gentilcore S, et al. Peer assessment of professional behaviours in problem-based learning groups. Med Educ. 2017;51(4):390–400. doi: 10.1111/medu.13151.
[71] Li H, Xiong Y, Hunter CV, et al. Does peer assessment promote student learning? A meta-analysis. Assess Eval High Educ. 2020;45(2):193–211. doi: 10.1080/02602938.2019.1620679.
[72] Deeb JG, Carrico CK, Laskin DM, et al. Influence of self-assessment on dental students' performance on technical assessments in periodontics. J Dent Educ. 2019;83(4):457–463. doi: 10.21815/JDE.019.047.
[73] Quick KK. The role of self- and peer assessment in dental students' reflective practice using standardized patient encounters. J Dent Educ. 2016;80(8):924–929. doi: 10.1002/j.0022-0337.2016.80.8.tb06172.x.
[74] Eva KW, Regehr G. "I'll never play professional football" and other fallacies of self-assessment. J Contin Educ Health Prof. 2008;28(1):14–19. doi: 10.1002/chp.150.
[75] Pintrich PR. The role of goal orientation in self-regulated learning. In: Handbook of self-regulation, research, and application. Orlando (FL): Academic Press; 2000. p. 451–502.
[76] Zimmerman BJ, Schunk DH. Self-regulated learning and academic achievement: theory, research, and practice. New York: Springer; 1989.
[77] Harris LR, Brown GTL. Using self-assessment to improve student Learning. 1st ed. Milton: Routledge; 2018.
[78] Panadero E, Andrade H, Brookhart S. Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. Aust Educ Res. 2018;45(1):13–31. doi: 10.1007/s13384-018-0258-y.
[79] Yan Z. Self-assessment in the process of self-regulated learning and its relationship with academic achievement. Assess Eval High Educ. 2020;45(2):224–238. doi: 10.1080/02602938.2019.1629390.
[80] Kalata LR, Abate MA. A mentor-based portfolio program to evaluate pharmacy students' self-assessment skills. Am J Pharm Educ. 2013;77(4):81–81. doi: 10.5688/ajpe77481.
[81] Wesely PM. Investigating the community of practice of world language educators on Twitter. J Teach Educ. 2013;64(4):305–318. doi: 10.1177/0022487113489032.
[82] Taras M. The use of tutor feedback and student self-assessment in summative assessment tasks: towards transparency for students and for tutors. Assess Eval High Educ. 2001;26(6):605–614. doi: 10.1080/02602930120093922.
[83] Lew MDN, Alwis WAM, Schmidt HG. Accuracy of students' self-assessment and their beliefs about its utility. Assess Eval High Educ. 2010;35(2):135–156. doi: 10.1080/02602930802687737.