



Using a theory-based approach to explore medical students' engagement with exemplar-based feedback and self-reflection for communication skills

Thea G. Musselman , Kendra Hester , Jason Snyder , Angelo D'Addario ,
Anastasiya A. Lipnevich & Krista Mattern

To cite this article: Thea G. Musselman , Kendra Hester , Jason Snyder , Angelo D'Addario ,
Anastasiya A. Lipnevich & Krista Mattern (07 Oct 2026): Using a theory-based approach to
explore medical students' engagement with exemplar-based feedback and self-reflection for
communication skills, Medical Teacher, DOI: [10.1080/0142159X.2026.2735289](https://doi.org/10.1080/0142159X.2026.2735289)

To link to this article: <https://doi.org/10.1080/0142159X.2026.2735289>

 View supplementary material 

 Published online: 07 Oct 2026.

 Submit your article to this journal 

 View related articles 

 View Crossmark data 

Using a theory-based approach to explore medical students' engagement with exemplar-based feedback and self-reflection for communication skills

Thea G. Musselman, Kendra Hester, Jason Snyder, Angelo D'Addario, Anastasiya A. Lipnevich and Krista Mattern

National Board of Medical Examiners, Philadelphia, Pennsylvania, USA

ABSTRACT

Introduction: Patient-Centered Communication (PCC) is essential to quality medical care and a crucial skill in medical education. While prior studies have focused on communication skills training in medical education, less is known about how learners engage in self-reflection when learning and practicing communication skills. The current study investigates levels of learner cognitive and reflective engagement with exemplar-based feedback and guided self-reflection to inform best practices for integrating self-regulated learning (SRL) supports in medical education.

Methods: Participants consisted of medical students in years one through three. Participants watched video vignettes depicting patients asking a question or making a statement, recorded verbal responses to the patients, reviewed their response alongside two exemplar responses, and completed a self-evaluation checklist aligned with defined communication learning objectives, and responded to three structured reflection prompts comparing their responses to the exemplars. Using an AI-assisted, human-in-the-loop coding process, participant engagement was evaluated using the ICAP taxonomy and Kember categorizations, each coded on a 1–4 scale reflecting framework-specific criteria. SRL strategies were coded as a separate unit of analysis that informed engagement.

Results: Participants demonstrated moderate to high engagement across the eight vignettes, characterized by predominantly *Active* (ICAP) and *Understanding* (Kember) level responses, with 91% of participants reaching peak *Constructive* (ICAP) engagement and 64% reaching peak *Reflective* (Kember) engagement at least once, along with widespread use of SRL strategies. Engagement varied between reflection prompts and across vignettes, with the third prompt and certain vignettes consistently eliciting deeper engagement.

Discussion: Results indicate that learners engaged meaningfully with exemplar-based feedback and self-reflection. Deeper engagement emerged with prompts that required comparison or improvement planning. While engagement varied by vignette and prompt, learners employed SRL strategies in nearly every reflection. These findings indicate that the use of exemplar responses and enhanced prompt design may support broader conceptual and deeper future learning.

ARTICLE HISTORY

Received 12 May 2026

Accepted 12 September 2026

KEYWORDS

Exemplar-based feedback; self-reflection; patient-centered communication

Effective patient-centered communication (PCC) is a keystone of high-quality healthcare. By supporting accurate diagnosis, fostering shared decision-making, building trust, and encouraging adherence to clinical recommendations, communication serves as a key mechanism through which care becomes safer, more effective, and more equitable (American College of Obstetricians and Gynecologists [ACOG], 2025; 1–7). Conversely, poor communication is associated with frequent adverse events, delays in treatment, medication errors, and even wrong-site surgeries [8,9]. Effective communication is also essential for advancing health equity. Substandard communication disproportionately impacts patients from historically marginalized communities, contributing to mistrust, reduced engagement, and unequal access to high-

quality care. By educating physicians to communicate with consideration of patients' cultural contexts, preferences, language needs, and social circumstances, patient-centered communication can help mitigate communication-related disparities and foster more equitable healthcare experiences and outcomes [10–13]. Moreover, quality communication does not only benefit patients. Physicians who engage in meaningful and robust communication with their patients report higher levels of job satisfaction, enabling a positive feedback loop that supports better patient care and a more fulfilling professional experience [14].

Consistent with its essential role in patient care, communication is emphasized internationally as a core competency for physicians. Educational frameworks across North America, Europe, the United

Practice points

- Learners engaged meaningfully with exemplar-based feedback, consistently demonstrating understanding and insight.
- Widespread use of self-regulated learning strategies shows that learners actively worked to comprehend, integrate, and apply feedback.
- Prompts that pushed learners beyond description toward reflection and planning elicited deeper engagement.
- Engagement variation across scenarios suggests that differences in complexity and interpersonal demands influence the depth of reflection.

Kingdom, and Australia explicitly identify interpersonal and communication skills as foundational to readiness for practice and require communication skills training as part of the curriculum [15–18].

Despite this emphasis, medical students often find developing communication skills challenging and report anxiety around patient communication when transitioning to clinical care [19–21]. Students cite several factors influencing their communication with patients, including specific training during medical school, personal identity, and experiences beyond medical education, and view communication as a nuanced process rather than the collection of discrete skills used for assessment in pre-clinical training [22]. Medical students report that they value communication skills training during medical school to provide structure for patient interactions and provide foundational knowledge of the skills expected in patient communication. However, much of their communication and interpersonal skills are learned through observation, and they question the utility of formal classroom instruction as it frequently conflicts with observation in real-world settings and note that feedback during clinical practice emphasizes medical content over communication [23–26].

Learning communication with feedback and reflection

While long thought of as an innate and immutable trait, both procedural and affective communication are teachable skills learned during medical education and that can improve with targeted communication education [27–34]. Reflection has been identified as an important component of communication skills development because it encourages learners to critically examine communication encounters, consider

patients' perspectives and emotions, and apply lessons learned to future interactions [35–38]. Reflective learning has consequently been linked to improved communication competence, empathy, and patient-centered care among medical learners [36,37,39].

Feedback is a key driver of improvement in communication performance [40–42] yet in real-world settings, personalized feedback is often constrained by faculty availability. To bridge this gap, promoting learners' ability to self-generate feedback through exemplars and guided reflection has emerged as a promising strategy [43–45]. Moreover, this approach positions learners as active participants in the learning process, fostering agency, autonomy, self-regulation, and receptivity to feedback; factors that have been linked to improved learning outcomes [46,47].

Exemplars, or high-quality examples of communication, provide concrete standards against which students can compare their own performance [48,49]. When combined with structured reflection prompts, exemplars encourage deeper engagement, helping learners identify strengths and areas for growth and formulate plans for improvement [50–52]. Existing research supports the efficacy of exemplar-based feedback and reflection for performance improvement, communication skill development, and related educational outcomes [35,36,47,49], but to our knowledge no studies have examined the cognitive, reflective, and self-regulatory processes medical students use while engaging with exemplar-based feedback and self-reflection in formative assessment and training in patient-centered communication. Understanding learner engagement is critical for refining interventions to maximize motivation, engagement, and ultimately, learning outcomes. This study centers on medical students' engagement with an intervention integrating exemplar-based feedback with guided self-reflection to patient-centered communication skills. Our goal is to provide insights into how learners interact with these instructional components, informing best practices for embedding self-regulated learning tools in medical education.

Defining engagement

We developed a multidimensional framework of engagement by utilizing Self-Regulated Learning (SRL) theory, the Kember Depth of Reflection categorizations, and the ICAP taxonomy. Given the complexity of the construct of engagement, we selected theories that encompass cognitive, metacognitive, and self-regulatory dimensions to represent complementary facets of engagement. SRL captures how learners intentionally direct and monitor their learning processes [53,54], ICAP characterizes the cognitive engagement processes learners employ while interacting

with instructional materials such as guided self-reflection activities [55], and Kember characterizes the depth of reflective interpretation demonstrated in written responses [56]. Prior research has linked self-regulation with deeper approaches to learning and reflection [57], suggesting that these constructs represent related but distinct aspects of learner engagement. Accordingly, we used the three frameworks as complementary analytic lenses to develop richer learner profiles and provide a more comprehensive understanding of learners' engagement with exemplar-based feedback and self-reflection.

Self-regulated learning

Self-regulated learning (SRL) describes how learners actively guide their own progress by setting goals, choosing strategies, monitoring their understanding, and adjusting their approaches as needed [53,54]. SRL engagement supports intentional, adaptive, and effective learning; it is strongly connected to deeper learning approaches [57] and can improve academic performance when supported with structured prompts [58]. In reflective writing, SRL occurs when learners identify goals for future performance, justify their strategy choices, and outline concrete plans for improvement. As part of the framework, we defined *Elaborating* as adding meaning to information by paraphrasing it in one's own words, *Rehearsing* as recalling information to maintain it in memory, and *Organizing* as structuring or linking ideas to show relationships and create a novel concept.

Kember depth of reflection

Kember categorizations outline four levels of depth in written reflections (Habitual Action, Understanding, Reflection, Critical Reflection) and describe how learners process information through writing. *Habitual Action* describes rote reporting of new information with little thought. *Understanding* occurs when learners absorb and apply content accurately but without deep analysis. *Reflection* demonstrates learners evaluating their performance, considering alternatives, and thinking critically about their choices. *Critical Reflection* describes when learners reassess their underlying assumptions, leading to meaningful improvements in how they approach future situations [56]. Research has shown that deeper forms of reflection are associated with more sophisticated integration of understanding [57].

ICAP taxonomy

The ICAP taxonomy (Interactive, Constructive, Active, Passive) explains how different levels of learner engagement contribute to different levels of learning. It distinguishes between simply receiving information (*Passive*),

manipulating or responding to it (*Active*), generating new ideas or inferences (*Constructive*), and co-constructing understanding with others (*Interactive*). These categories illustrate the difference between merely completing an activity versus engaging in the deeper cognitive work needed to build new understanding [55,59].

Together, these models capture a nuanced understanding of engagement by encompassing how learners guide their own learning (SRL), how deeply they reflect (Kember) and how they process information (ICAP). Instead of viewing the three theories as independent constructs, we frame them as complementary elements of the engagement process. Examining all three dimensions simultaneously provides a more comprehensive understanding of learner engagement than any framework alone.

The current study

This research is part of a larger research study examining whether integrating exemplar-based feedback, self-assessment, and structured self-reflection within the Communication Learning Assessment (CLA), a tool designed to support the development of PCC skills, improves students' ability and confidence in applying these skills. Participants were assigned to one of four conditions that varied in access to self-directed feedback resources including exemplar responses, self-evaluation checklists, and guided self-reflection prompts. This paper reports on the participant condition that had access to all aspects of guided self-reflection, focusing on depth of engagement with the activities, while the experimental results from the main study have been reported elsewhere [60].

The specific aim of this study is to examine depth of medical students' cognitive engagement and reflective engagement with exemplar-based feedback and guided self-reflection to understand how learners interact with these instructional components and inform best practices for embedding self-regulated learning supports in medical education.

Methods

Recruitment

Participants were recruited through medical school postings, USMLE announcements, and outreach to Executive Chief Proctors. Eligibility was limited to medical students, who were 1) age 18 and above, 2) attending a US-based school, 3) and had not yet completed their first clerkship rotation. All participants provided informed consent under an IRB-approved protocol (EX00714), and confidentiality was maintained through anonymization of all data.

Feedback Activity

You are speaking with Rick, the son of John Wilson, who is currently hospitalized for abdominal pain and fever. His fasting ultrasound was canceled after Rick brought and fed him breakfast. When Rick learned the test was canceled, he demanded to speak to you. You enter the room and introduce yourself, and Rick says:

Below are the aspects of patient-centered communication that are present in the exemplar responses. In this self-reflection exercise, check those aspects of communication that are present in your response and describe how your response compares to the two exemplar responses.

▶Your Response ▶Exemplar 1 ▶Exemplar 2

Your Response	Exemplar 1	Exemplar 2	Patient-Centered Communication Behavior (PCB)
☐	☒	☒	Take responsibility and apologize for the miscommunication.
☐	☐	☒	Provide validation of patient's emotion.
☐	☒	☐	Assure the patient that the procedure will be promptly rescheduled.

In what ways is your response similar to the exemplar responses?

Type your response here

In what ways is your response different from the exemplar responses?

Type your response here

What are you planning on changing to improve your patient-centered communication in future encounters?

Type your response here

Figure 1. Feedback and reflection Activity.

Procedure

Participants responded to eight video-based patient communication vignettes across two sessions. The vignettes were developed by clinicians and medical educators using an iterative process that included pre-testing each vignette and reviewing them for realism and their ability to elicit meaningful communication behaviors [61]. Exemplar responses were selected from recordings previously completed by learners during residency that received high ratings from lay patient raters and contained the targeted patient-centered communication behaviors associated with each vignette, allowing them to serve as models of effective communication. For each vignette, students first reviewed background information about the clinical scenario, watched a brief video of a patient asking a question or making a statement, and recorded a verbal response as if speaking directly to the patient. After recording their initial response, participants listened to their own response and reviewed two exemplar responses demonstrating targeted patient-centered communication behaviors (see Supplemental Table 1 for vignette descriptions, PCBs, and exemplars). They

then completed a self-evaluation checklist aligned with the communication objectives and responded to three reflection prompts asking how their response was similar to the exemplars, how it differed, and what changes they would make in future patient interactions (Figure 1). Finally, participants were given an opportunity to re-record their response to the same vignette. This sequence was repeated for each of the eight vignettes.

Coding framework

We developed a coding framework guided by processes described in the literature [62–64], using the complementary theoretical models of SRL, Kember Depth of Reflection categorizations, and the ICAP taxonomy. To develop an applied inductive codebook, two analysts each reviewed 15% of unique responses, developed initial coding definitions, and synthesized results to operationalize codes. The three emergent SRL strategies were *Elaborating* by putting a PCB in one's own words, *Rehearsing* by describing a patient communication strategy from an exemplar that was not a defined PCB, and *Organizing* by linking information to create a novel

Table 1. SRL strategies, Kember, and ICAP coding categories.

Coding classification	Definition	Indicators
Self-regulated learning strategies		
Organizing: <i>linking information</i>	Structuring or linking ideas to show relationships and create a coherent concept.	Learner links ideas and previous experiences to describe a PCC behavior that is not part of the provided PCBs or exemplars.
Elaborating: <i>paraphrasing</i>	Adding meaning to information by paraphrasing it in one's own words.	Learner paraphrases a PCB in their own words.
Rehearsing: <i>retrieving and restating</i>	Repeating or recalling information to maintain it in memory.	Learner references a communication strategy present in exemplar that is not a provided PCB.
Rote Repetition	Surface-level behavior signaling low engagement.	Learner restates provided PCB text without rewording, interpreting, or manipulating.
Kember levels of reflective depth		
Critical Reflection (L4)	Evidence of a change in perspective over a fundamental belief; requires transformation of presuppositions; rare.	Explicit perspective shift, reassessing underlying assumptions, transformative re-interpretation of experience. Indicated by "I questioned/I recognized/ I reconsidered" language. Typically evident in longer responses.
Reflection (L3)	Applying theory to practical situations; integrating personal experiences with concepts; generating personal insights beyond book theory.	Clear linkage of PCB/exemplar concepts to the learner's own understanding; insights about why/how their approach works or needs to change. Indicated by <i>Organizing</i> and "why/how" language.
Understanding (L2)	Correct explanation of theory without connection to personal experience; conceptual but not applied.	Accurate description of PCBs/exemplar ideas without personal relevance or applied reasoning. Indicated by use of <i>Elaborating</i> or <i>Rehearsing</i> and/or "I know that/ this relates to/I understand" language.
Habitual Action (L1)	Non-reflection: routine, procedural, or reproduced content without interpretation.	Repeating material, generic statements; no reasoning or personal meaning. Indicated by <i>Rote Repetition</i> or no substantive content.
ICAP cognitive engagement categories		
Interactive (L4)	Dialoguing: both partners generate constructive contributions; co-inference of new ideas.	Integrating multiple viewpoints, iterative reasoning. Not common in point-in-time activities. Indicated by change over time or upon external direct feedback.
Constructive (L3)	Generating ideas beyond the provided information: self-explaining, inferring, integrating, linking concepts.	Use of critical reasoning, novel linkage, generative explanation. Indicated by <i>Organizing</i> and "how/why" language.
Active (L2)	Manipulating information without adding meaning: highlighting, listing, selecting, identifying.	Describes actions, strategies, or information provided but without inference, evaluation, or causal reasoning. Indicated by <i>Elaborating</i> or <i>Rehearsing</i> .
Passive (L1)	Receiving information without additional processing.	Repeating information exactly as written on feedback screen. Indicated by <i>Rote Repetition</i> and response with no substantive content.

concept [54,65]. We also coded *Rote Repetition* under the SRL heading for verbatim PCB mentions, which we interpreted as demonstrating surface-level engagement.

ICAP and Kember classifications were informed by SRL strategy use and interpretation of the literature. Using the Kember framework [56], reflective depth was classified as *Habitual Action*, *Understanding*, *Reflection*, or *Critical Reflection*. Using the ICAP taxonomy [55], cognitive engagement was classified as *Passive*, *Active*, *Constructive*, or *Interactive*. See Table 1 for definitions and coding indicators.

AI-assisted coding and HITL process

We used GPT-5 LLM accessed through Co-Pilot with a human-in-the-loop (HITL) workflow for coding, where humans consistently reviewed human versus AI agreement to refine prompting instructions and LLM classification. Two analysts independently double-coded 25% of responses for SRL strategies and met to resolve discrepancies and establish consensus. Following this calibration process, the remaining responses were divided evenly and coded

independently for SRL strategies using Atlas.ti. The 25% training set was also human-coded for ICAP and Kember levels to establish a ground truth dataset for assessing human–LLM agreement.

Reflections, human-coded SRL strategies, original academic articles describing Kember and ICAP, supplementary coding logic, a project description and analysis plan, and an output template were provided as input to the LLM, which generated ICAP and Kember codes. Agreement between the LLM outputs and the human-coded ground truth set was evaluated using Cohen's kappa. Aligned with commonly used classifications of Cohen's kappa, a threshold of 0.80 was used to indicate substantial or near-perfect agreement [66,67]. Cohen's kappa was calculated by comparing human and LLM binary present/absent classifications across vignettes for ICAP, Kember, SRL strategies, exemplars, and PCBs. When agreement fell below 0.80 for any vignette, the coding logic and prompting guidelines were revised and the LLM-coding was re-run. Once acceptable agreement was achieved for all vignettes, human coders performed quality assurance on all output, including reviewing cases where ICAP and Kember levels

Table 2. ICAP By reflective prompt.

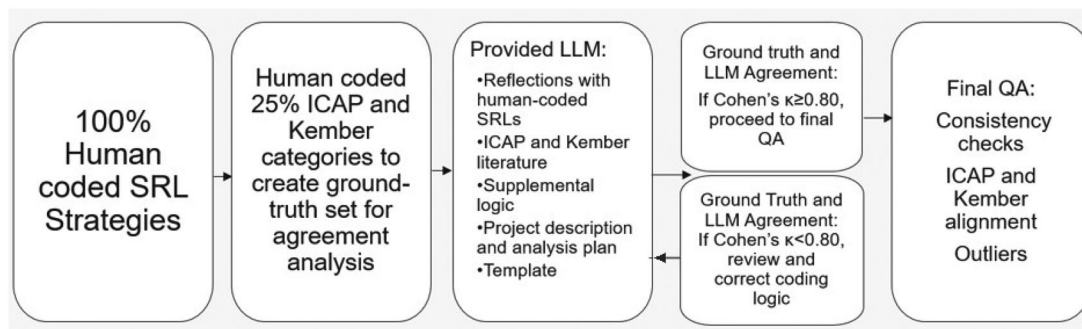
Reflective prompt	Passive: <i>N</i> (%)	Active: <i>N</i> (%)	Constructive: <i>N</i> (%)	<i>M</i> (SD)	<i>M</i> (SD) of peak
Reflective Prompt 1 <i>Similarities</i>	121 (35%)	216 (62%)	11 (3%)	1.68 (.29)	2.16 (0.52)
Reflective Prompt 2 <i>Differences</i>	123 (35%)	201 (58%)	24 (7%)	1.71 (.30)	2.33 (0.48)
Reflective Prompt 3 <i>Future Plans</i>	82 (24%)	156 (45%)	110 (32%)	2.07 (.38)	2.89 (0.38)

Note: ICAP level categories (passive, active, and constructive) represent the distribution of responses ($n = 1044$) for each reflective prompt. Mean (standard deviation) represents the learner-level ($n = 45$) average and then averaged across learners. Peak represents the highest score achieved per learner, summarized across learners.

Table 3. Kember by reflective prompt.

Reflective prompt	Habitual: <i>N</i> (%)	Understanding: <i>N</i> (%)	Reflection: <i>N</i> (%)	<i>M</i> (SD)	<i>M</i> (SD) of peak
Reflective Prompt 1 <i>Similarities</i>	101 (29%)	247 (71%)	0 (0%)	1.71 (.24)	1.96 (0.21)
Reflective Prompt 2 <i>Differences</i>	112 (32%)	220 (63%)	16 (5%)	1.72 (.24)	2.27 (0.45)
Reflective Prompt 3 <i>Future Plans</i>	83 (24%)	195 (56%)	70 (20%)	1.96 (.38)	2.56 (0.55)

Note: Kember level categories (habitual, understanding, and reflection) represent the distribution of responses ($n = 1044$) for each reflective prompt. Mean (standard deviation) represents the learner-level ($n = 45$) average and then averaged across learners. Peak represents the highest score achieved per learner, summarized across learners.

**Figure 2.** HITL LLM coding process.

appeared discordant to ensure logical consistency across engagement and reflection classifications. See Figure 2.

Analytical approach

We assigned 1–4 score values from lowest to highest levels of engagement for ICAP and Kember categories and analyzed those data at the overall, reflective prompt, and vignette levels. For reflective prompt-level analyses, we averaged each of the learner's eight scores for a given reflective prompt across vignettes to produce a learner-level mean, then averaged these means across learners to generate reflective prompt-level means and standard deviations, along with the distribution of reflections across the four ICAP and Kember categories. Vignette-level summaries followed a parallel procedure, where we averaged each learner's three scores within a vignette (one per reflective prompt). We calculated overall scores by averaging all 24 reflections for each learner, then averaging these learner-level means across the full sample, similarly reporting accompanying category frequencies.

To characterize learners' highest demonstrated levels of engagement and reflection, we determined peak (maximum) ICAP and Kember scores at each analytic level. We identified an overall peak score for each learner by selecting the highest score across all

24 reflections. For each of the three reflective prompts, we identified the maximum score for that reflective prompt a learner achieved across the eight vignettes and summarized these maxima across learners. For each vignette, we identified the highest score the learner achieved across the three reflective prompts and summarized the resulting distribution. At each level, we report the mean and standard deviation of these maximum values, as well as the percentage of learners whose peak fell within each ICAP and Kember category.

We operationalized SRL strategy use at each analytic level (overall, reflective prompt, and vignette) by coding each strategy as a binary learner-level indicator (used vs. not used) and then reporting the proportion of learners who used a given strategy at least once within that unit of analysis:

- Overall level: If it appeared in any of their 24 reflections, we coded a learner as having used a strategy.
- Reflective prompt level: If it appeared in any of their eight associated reflections, we coded a learner as having used a strategy.
- Vignette level: If it appeared in any of their three reflections, we coded a learner as having used a strategy.

The resulting counts and percentages therefore represent the binary prevalence of strategy use

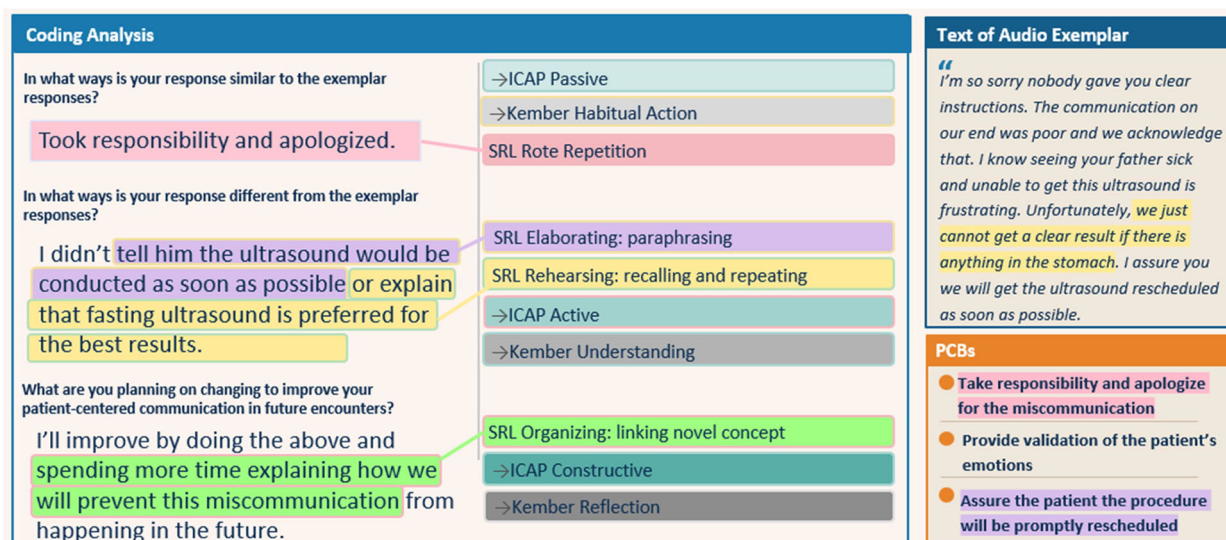


Figure 3. Illustrative responses with applied coding.

among learners at each unit of analysis, independent of how frequently individual learners used a given strategy.

We summarized word count using parallel procedures. For each learner, we calculated the mean and standard deviation of word count for the eight reflections to each reflective prompt, for the three reflections within each vignette, and for all 24 reflections combined. We then averaged learner-level means and standard deviations across the sample for each level of analysis.

Results

Participants

Participants were 45 students in their first-year ($n=3$; 7%), second-year ($n=28$; 62%), or third-year ($n=14$; 31%) of medical school, and few ($n=7$; 16%) had entered their first clerkship. The group included female ($n=34$; 76%), male ($n=10$; 22%) and transgender ($n=1$; 2%) students. Racial and ethnic identities included learners identifying as White ($n=15$; 33%), Black ($n=10$; 22%), Asian ($n=10$; 22%), Hispanic or Latino ($n=3$; 7%), and Multiracial ($n=5$; 11%), with 4% ($n=2$) not reporting.

Reflection characteristics

The data set includes a total of 1044 reflection responses, with between 42 and 45 responses per vignette. Although participants were required to enter text into the self-reflection field to proceed to the next vignette, a technical error with the platform prevented a small subset of responses from saving ($n=36$; 3.3%). Learners produced an average of 15.88 words per reflection ($SD=13.3$) with lengths ranging from 1 to 129 words.

Overall engagement

Across the full assessment (3 reflections $\times$ 8 vignettes), learners demonstrated consistently high engagement with the self-reflection tasks. SRL behaviors were nearly ubiquitous: 100% of learners engaged in *Organizing*, 100% engaged in *Elaborating*, and 98% engaged in *Rehearsing*, showing that learners consistently employed multiple SRL strategies over the course of the eight-vignette assessment. Figure 3 provides illustrative responses with applied codes.

In terms of Kember frequencies across all reflections, the majority of responses (65%) were classified as *Understanding* (L2), and the overall Kember mean was 1.80 ($SD=0.57$). The majority of learners (64%) reached *Reflection* (L3) as their peak level, and the participant-level peak Kember mean reached 2.64 ($SD=0.48$).

Similarly, in terms of ICAP frequencies across all reflections, learners most often demonstrated *Active* (L2) engagement (55% of all responses), followed by *Passive* (L1) engagement (31% of all responses), with *Constructive* (L3) responses making up 14%. Mean overall ICAP engagement score was 1.83 ($SD=0.65$). For most participants (91%), their peak ICAP engagement level was *Constructive* (L3). On average, the peak ICAP score across all responses per learner was 2.91 ($SD=0.29$).

Reflective prompt engagement

Across the eight vignettes, the three reflective prompts elicited distinct patterns of engagement. Looking at SRL behaviors, learners demonstrated the lowest rates of *Organizing* on reflective prompt 1 (How was your response similar?; 22 learners; 48.9%), a marked increase on reflective prompt 2 (How was your response different?; 34 learners; 75.6%), and the

Table 4. Kember by vignette.

Vignette	N (participants/ responses)	Habitual: N (%)	Understanding: N (%)	Reflection: N (%)	M (SD)	M (SD) of peak
Unneeded testing request	43 (129)	23 (18%)	93 (72%)	13 (10%)	1.92 (0.32)	2.28 (0.45)
End of life options	43 (129)	28 (22%)	90 (70%)	11 (8%)	1.87 (0.38)	2.16 (0.53)
Infectious disease precautions	45 (135)	30 (22%)	93 (69%)	12 (9%)	1.87 (0.37)	2.20 (0.50)
Approaching sensitive topic	43 (129)	40 (31%)	75 (58%)	14 (11%)	1.80 (0.41)	2.23 (0.57)
Miscommunication	41 (123)	31 (25%)	88 (72%)	4 (3%)	1.78 (0.37)	1.98 (0.47)
Managing patient frustration	45 (135)	40 (30%)	88 (65%)	7 (5%)	1.76 (0.36)	2.09 (0.47)
Clarify medication needs	43 (129)	47 (36%)	71 (55%)	11 (9%)	1.72 (0.40)	2.12 (0.63)
Explaining new diagnosis	45 (135)	57 (42%)	64 (47%)	14 (10%)	1.68 (0.54)	1.98 (0.78)

Note: Kember level categories (habitual, understanding, and reflection) represent the distribution of responses for each vignette. Mean (standard deviation) represents the learner-level average within a vignette and then averaged across learners. Peak represents the highest score achieved per learner for a vignette, summarized across learners.

highest rates on reflective prompt 3 (What will you change in future encounters?; 44 learners; 97.8%). *Elaborating* was consistently high across prompts, occurring in 93.3% of responses on reflective prompt 1, 97.8% on reflective prompt 2, and 91.1% on reflective prompt 3. Rehearsing was most prevalent on reflective prompt 1 (88.9%) and reflective prompt 2 (95.6%), before declining on reflective prompt 3 (80.0%). In contrast, Rote Repetition showed a steady decrease across reflective prompts, from 68.9% on reflective prompt 1, to 60.0% on reflective prompt 2, and 31.1% on reflective prompt 3.

Learners demonstrated the lowest Kember levels on reflective prompt 1 ($M=1.71$, $SD=0.24$), with no responses classified as *Reflection* (L3), followed by a modest increase on reflective prompt 2 ($M=1.72$, $SD=0.24$) with 5.0% of responses reaching the *Reflection* (L3) level. The highest reflective depth was observed on reflective prompt 3 ($M=1.96$, $SD=0.38$), with 20.0% of responses classified as *Reflection* (L3). Participant-level peak Kember scores showed the same pattern, with reflective prompt 3 yielding the highest mean peak score ($M=2.56$, $SD=0.55$), and over half of participants (58.0%) reaching *Reflection* (L3) at least once.

Learners demonstrated the lowest ICAP engagement on reflective prompt 1 ($M=1.68$, $SD=0.29$), with only 3.2% of responses classified as *Constructive* (L3), followed by a modest increase on reflective prompt 2 ($M=1.71$, $SD=0.30$), with 7.0% of responses classified as *Constructive* (L3). The highest engagement occurred on reflective prompt 3 ($M=2.07$, $SD=0.39$), where 32.0% of responses were classified as *Constructive* (L3). Participant-level peak ICAP values followed the same progression, with reflective prompt 3 yielding the highest mean peak score ($M=2.89$, $SD=0.38$), and nearly all learners (91.0%) reaching the *Constructive* (L3) level at least once.

Vignette engagement

SRL strategy use varied across vignettes. The number of learners who engaged in *Organizing* was highest in the approaching Sensitive Topic vignette (29;

67.4%) and the Infectious Disease Precaution vignette (28; 62.2%), and lowest in the Explaining New Diagnosis vignette (7; 15.6%). *Elaborating* was consistently common across all vignettes, but with particularly lower rates in the Explaining New Diagnosis vignette (25; 55.6%). *Rehearsing* showed substantial variability, ranging from 12 (27.9%) in the Approaching Sensitive Topic vignette to 37 (86.0%) in the Unneeded Testing Request vignette. *Rote Repetition* occurred most frequently in Managing Patient Frustration (22; 48.9%) and least frequently in the Infectious Disease Precaution vignette (7; 15.6%).

Kember means demonstrated the highest levels of reflective depth for the Unneeded Testing Request vignette ($M=1.92$, $SD=0.32$), the End of Life Options vignette ($M=1.87$, $SD=0.38$), and the Infectious Disease Precaution vignette ($M=1.87$, $SD=0.37$). Across these vignettes, *Reflection* (L3) occurred in 10%, 8%, and 9% of responses, respectively. These vignettes also had the lowest frequency of *Habitual Action* (L1) classifications, occurring 18%, 22%, and 22% of responses, respectively. Peak Kember scores were especially high for the Unneeded Testing Request vignette ($M=2.28$, $SD=0.45$), the Approaching Sensitive Topic vignette ($M=2.23$, $SD=0.57$), and the Infectious Disease Precaution vignette ($M=2.20$, $SD=0.50$) with at least 25% of participants reaching *Reflection* (L3) and only 7% peaking at *Habitual Action* (L1) (refer to Table 4 for more detailed results). Responses to the Unneeded Testing Request vignette showed higher levels of reflective engagement earlier in the sequence of reflective prompts and sustained engagement. For example, Prompt 1 evoked the *Active* (L2) response "This response is similar such that I acknowledge the patient and validate the discomfort which are key components to patient centered communication before delving into the reasoning for avoiding imaging," with the participant using *Rehearsing* by recalling communication strategies used only in the exemplars and *Elaborating* by paraphrasing and rewording given PCBs more frequently than the average across vignettes. Similarly, this response to Prompt 2, "I realized I need to make sure to be more explicit about what I am

Table 5. ICAP by vignette.

Vignette	N (participants/ responses)	Passive: N (%)	Active: N (%)	Constructive: N (%)	M (SD)	M (SD) of peak
Approaching sensitive topic	43 (129)	29 (22%)	73 (57%)	27 (21%)	1.98 (0.39)	2.56 (0.50)
Unneeded testing request	43 (129)	26 (20%)	83 (64%)	20 (16%)	1.95 (0.35)	2.42 (0.50)
Infectious disease precautions	45 (135)	34 (25%)	74 (55%)	27 (20%)	1.95 (0.45)	2.42 (0.62)
Miscommunication	41 (123)	35 (28%)	68 (55%)	20 (16%)	1.88 (0.50)	2.22 (0.72)
Clarify medication needs	43 (129)	51 (40%)	57 (44%)	21 (16%)	1.77 (0.48)	2.26 (0.76)
End of life options	43 (129)	39 (30%)	81 (63%)	9 (7%)	1.77 (0.37)	2.12 (0.50)
Managing patient frustration	45 (135)	49 (36%)	69 (51%)	17 (13%)	1.76 (0.42)	2.22 (0.64)
Explaining new diagnosis	45 (135)	63 (47%)	68 (50%)	4 (3%)	1.56 (0.47)	1.76 (0.61)

Note: ICAP level categories (passive, active, and constructive) represent the distribution of responses for each vignette. Mean (standard deviation) represents the learner-level average within a vignette and then averaged across learners. Peak represents the highest score achieved per learner for a vignette, summarized across learners.

recommending and provide information on why I am not recommending something if applicable in order to ensure patient understanding” demonstrated *Reflection (L3)* as the learner expressed why changing their communication approach is necessary whereas other vignettes typically did not elicit *Reflection (L3)* until prompt 3.

Reflective depth was lowest for the Explaining New Diagnosis vignette ($M = 1.68$, $SD = 0.54$), which had the highest frequency of *Habitual Action (L1)* (42% of responses). Responses to this vignette were characterized by relatively superficial responses across the three prompts. For example, the response “Check for understanding” utilized only *Rote Repetition* by stating a PCB verbatim.

Mean ICAP scores ranged from 1.56 to 1.98, with the highest levels observed for the Approaching Sensitive Topic vignette ($M = 1.98$, $SD = 0.39$) with 21% of responses classified as *Constructive (L3)*, Unneeded Testing Request ($M = 1.95$, $SD = 0.35$) with 16% of responses classified as *Constructive (L3)*, and the Infectious Disease Precaution vignette ($M = 1.95$, $SD = 0.45$) with 21% of responses classified as *Constructive (L3)*. Participant-level peak ICAP values aligned with these trends, again showing the strongest upper end performance in the Approaching Sensitive Topic vignette ($M = 2.56$, $SD = 0.50$) with 55% of participants reaching *Constructive (L3)*, the Unneeded Testing Request vignette ($M = 2.42$, $SD = 0.50$) with 42% of participants reaching *Constructive (L3)*, and the Infectious Disease Precaution vignette ($M = 2.42$, $SD = 0.62$) with 49% of participants reaching *Constructive (L3)*. Again, learners’ responses to these vignettes elicited higher levels of engagement earlier in the feedback reflective prompt sequence and demonstrated sustained engagement. For example, this response to reflective prompt 2 “I didn’t specifically [address the topic]. I was worried about sounding accusatory, but I realize it is important to make the patient aware of health risks so that she has that information available and can make informed decisions” showed reasoning and evaluative thinking in a way that is more frequently characterized by responses to reflective prompt 3.

Similarly to the Kember results, the lowest engagement appeared for the Explaining New Diagnosis vignette ($M = 1.56$, $SD = 0.47$) with only 3% of responses classified as *Constructive (L3)*, and the lowest peak engagement for the Explaining New Diagnosis vignette ($M = 1.76$, $SD = 0.61$) with only 9% of participants reaching *Constructive (L3)* (refer to Table 5 for more detailed results).

Discussion

While prior studies have largely focused on learning outcomes and performance improvement [35,36,47,49], this study examined the engagement processes underlying exemplar-based feedback and reflection. By examining engagement through SRL, ICAP, and Kember frameworks, the study provides insight into the cognitive, reflective, and self-regulatory processes learners use while interacting with exemplar-based feedback and structured reflection.

The results demonstrate that across vignettes and reflection prompts, learners consistently demonstrated behaviors that indicated they were absorbing, processing, and accurately interpreting the communication concepts presented in the PCBs and exemplars. Learners uniformly employed at least one SRL strategy in nearly every reflection, indicating deeper levels of cognitive processing. Reflections dominated by *Active (L2)* ICAP classifications (55%) and *Understanding (L2)* level responses on the Kember scale (65%) illustrates that most learners didn’t simply repeat the PCBs listed in the checklist but rather correctly identified, described, and interpreted concepts in ways that reflected comprehension. These results, along with the high rates of learners reaching *Constructive (L3)* or *Reflection (L3)* levels at least once, indicate that students reliably absorbed PCC concepts and articulated their meaning accurately, even if not yet displaying the highest levels of engagement.

These data support the utility of exemplars as tools for feedback that extends beyond simply being a tool to demonstrate clear performance standards to learners. Across reflective prompts and vignettes, learners frequently referenced PCC behaviors that

were not vignette-specific provided PCBs but which were present in the exemplars. This suggests that exemplars can be sources of new information in addition to providing a stimulus for self-directed feedback and self-reflection. High rates of *Rehearsing* (98%) show that learners recognized and interpreted exemplar concepts, which reflects students' ability to link exemplar illustrations with PCC skills beyond the defined PCBs, synthesizing information in novel ways. This indicates that exemplars may serve as more than illustrations of specific PCBs, but as unique sources of new concepts for learners to extract and interpret. Additionally, this trend demonstrates deep engagement with the exemplars and highlights the potential value of exemplars for fostering learning. Differences across the three reflective prompts illustrate that prompts theoretically aligned with SRL models may effectively evoke deeper engagement. The observed progression in engagement depth found across the three reflective prompts aligns with literature that posits prompt structure and cognitive demands strongly influence the depth of learner engagement [68–70]. Simply asking learners to identify similarities (reflective prompt 1) elicited mostly surface-level engagement, reflecting what Jasper [70] describes as descriptive rather than analytical reflection. Alternately, prompts asking learners to identify differences (reflective prompt 2) and plan specific changes (reflective prompt 3) reliably produced higher ICAP and Kember levels, mirroring evidence showing reflective prompts that require evaluation, contrast, and planning more efficiently facilitate self-regulated learning [68]. Reflective Prompt 3 explicitly prompted learners to engage in the SRL strategy of articulating how they would change their future communication, thus reflective prompt 3 elicited the highest rates of *Constructive* (L3) engagement and *Reflection* (L3). However, learners may benefit from a scaffolding effect where earlier reflective prompts help students build foundational understanding before reaching deeper levels [68,71,72], implying the utility of the reflective prompt 1 and reflective prompt 2, despite lower levels of engagement demonstrated compared to reflective prompt 3.

Educational theory suggests that learners engage most deeply when tasks present challenges that are neither too simple nor overly complex, a state often described as productive struggle. Productive struggle refers to effortful engagement with challenges just beyond a learner's current abilities within the zone of proximal development [73–76]. In the present study, engagement varied across vignettes, suggesting that characteristics of the cases themselves influenced the depth of learners' reflection. Although causal mechanisms cannot be determined from this

single study, one possible explanation is that learners differed in their familiarity and comfort with the clinical and interpersonal demands of each vignette, which may have positioned some cases as more optimally challenging than others. Learners demonstrated most shallow engagement with the *Explaining New Diagnosis* vignette, which centered on providing clinical knowledge gained early in medical education while placing a low interpersonal demand on the student. The relative ease of this scenario may have provided insufficient productive struggle for the learner to engage deeply. Conversely, the vignette focused on approaching a sensitive topic and the one focused on refusing an unneeded testing request involved a higher level of interpersonal tension, which elevates the difficulty of the situation and thus may elicit deeper engagement. For familiar clinical scenarios, engagement may be increased by adjusting the clinical context and interpersonal demand to an optimal level of challenge for the learner level. Other factors impacting engagement may include patient affect and characteristics, which has been shown to impact how physicians communicate with patients and thus may impact their reflective engagement [77–79]. Identifying which vignette features support consistent, high-quality engagement will be important for designing diverse content while maintaining opportunities for meaningful cognitive and reflective engagement.

Unsurprisingly, no reflections reached the ICAP *Interactive* (L4) level or Kember *Critical Reflection* (L4) levels, which was expected given the brief length of responses and the self-guided, point in time nature of the reflections. Reaching the highest levels typically requires dialogue and iterative feedback, which is not a component of this self-directed exercise. However, learners' self-reflections can focus their discussions with faculty to evoke more targeted feedback that promotes the collaborative iteration and growth needed to reach the *Interactive* (L4) level. Although the activities may not evoke these peak levels, the reflections can serve as a foundation to help learners progress toward the higher levels of engagement.

Taken together, these findings have several implications for communication skills education. Reflection prompts that require learners to identify specific future behavior changes demonstrated deeper engagement than prompts focused on simple comparisons, suggesting that action-oriented reflection may be especially valuable in communication training. The findings also support the use of exemplar responses as both feedback and instructional tools, helping learners identify communication strategies and generate plans for improvement.

These results suggest that communication curricula may benefit from integrating exemplars, structured self-evaluation, and future-focused reflection prompts to promote self-regulated learning and engagement.

Limitations

This study has several limitations. Low variability in ICAP and Kember classifications, with responses falling between Levels 1–3, as well as limited variability in performance on CLA for this study condition and the small sample size, limited our ability to examine the relationship between engagement depth and demonstrated PCC skill. Future studies should include larger and more diverse samples, as well as conditions that elicit a broader range of performance and engagement, to better capture these relationships. In addition, as described in the Results, a technical issue in the platform resulted in missing reflection data for a subset of responses. Although the amount of missing data was small (3.3%), it may have affected the distribution of engagement levels for certain vignettes or reflective prompts. Future work should ensure more robust data capture and monitoring procedures to minimize data loss and support more complete analyses. Finally, the brevity of the written responses may have limited the detection of reflective processes and likely contributed to the absence of higher-level classifications (*Interactive* or *Critical Reflection*).

Another limitation is the potential influence of prior clinical exposure and learning through observation on engagement with the reflective activities. Although only a small number of participants had begun clerkship training and none had completed a clinical rotation, exposure to clinical environments shape learners' perceptions of communication and reflection [26,80]. Because the sample was too small to examine differences by training level, future research should explore how learning through observation influences engagement, reflective depth, and communication skill development.

Conclusions

This study demonstrates that exemplar-based feedback combined with structured self-reflection can elicit meaningful cognitive and reflective engagement in medical students learning PCC skills. Reflection prompt design played a key role in supporting deeper engagement, with reflection questions requiring contrasting one's own performance with exemplars and planning future behavior modifications eliciting the highest levels of engagement. Engagement depth was also influenced by the

vignette characteristics, suggesting that instructional stimuli interact with reflective demands to shape learners' engagement. Together, these findings highlight the value of intentionally aligning reflection scaffolding and vignette design to support self-regulated learning in communication training.

Author contributions

CRedit: **Thea G. Musselman**: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing; **Kendra Hester**: Formal analysis, Visualization, Data curation, Writing – original draft, Writing – review & editing; **Jason Snyder**: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing; **Angelo D'Addario**: Conceptualization, Data curation, Formal analysis, Project administration, Supervision, Writing – review & editing; **Anastasiya A. Lipnevich**: Investigation, Project administration, Supervision, Writing – review & editing; **Krista Mattern**: Conceptualization, Investigation, Project administration, Supervision, Writing – review & editing.

Disclosure statement

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

Funding

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

Notes on contributors

Thea G. Musselman, MPH; Research Associate, Office of Research Strategy, National Board of Medical Examiners.

Kendra Hester, MPH; Research Associate, Office of Research Strategy, National Board of Medical Examiners.

Jason Snyder, MA; Research Associate, Office of Research Strategy, National Board of Medical Examiners.

Angelo D'Addario, MS; Research Associate Manager, Office of Research Strategy, National Board of Medical Examiners.

Anastasiya A. Lipnevich, PhD; Principal Measurement Scientist, Office of Research Strategy, National Board of Medical Examiners.

Krista Mattern, PhD; Vice President, Office of Research Strategy, National Board of Medical Examiners

References

1. Arbuthnott A, Sharpe D. The effect of physician–patient collaboration on patient adherence in non-psychiatric medicine. *Patient Educ Couns*. 2009;77(1): 60–67. doi:10.1016/j.pec.2009.03.022
2. Bukstein DA. Patient adherence and effective communication. *Ann Allergy Asthma Immunol*. 2016; 117(6):613–619. doi:10.1016/j.anai.2016.08.029

3. Butow P, Sharpe L. The impact of communication on adherence in pain management. *Pain*. 2013;154 Suppl 1(Supplement 1): s 101–S107. doi:10.1016/j.pain.2013.07.048
4. Howick J, Moscrop A, Mebius A, et al. Effects of empathic and positive communication in healthcare consultations: a systematic review and meta-analysis. *J R Soc Med*. 2018;111(7):240–252. doi:10.1177/0141076818769477
5. Lu Chen L, Chu H, Wang H, et al. Effects of communication, trust, and respect on shared decision-making. *Front Med (Lausanne)*. 2025;12:1577276. doi:10.3389/fmed.2025.1577276
6. Song M, Elson J, Haas C, et al. The effects of patients' health information behaviors on shared decision-making. *Healthcare*. 2025;13(11):1238. doi:10.3390/healthcare13111238
7. Zolnieriek KBH, DiMatteo MR. Physician communication and patient adherence: a meta-analysis. *Med Care*. 2009;47(8):826–834.
8. Howick J, Bennett-Weston A, Solomon J, et al. How does communication affect patient safety? Protocol for a systematic review and logic model. *BMJ Open*. 2024;14(5): E085312. doi:10.1136/bmjopen-2024-085312
9. Shitu Z, Hassan I, Aung MMT, et al. Avoiding medication errors through effective communication. *Malays J Mov Health Exerc*. 2018;7(1):115–128. doi:10.4103/2600-9404.323043
10. Aidoo EAK, Wood SF, Issaka B, et al. Patient-centered communication and its impact on minority health competence and outcomes in the United States. *Sci Rep*. 2025;15(1):11885. doi:10.1038/s41598-025-97536-9
11. Leotin S. Addressing health communication gaps: improving patient experiences and outcomes through human-centered design. *J Patient Exp*. 2025;12:23743735251334015. doi:10.1177/23743735251334015
12. Stubbe DE. Practicing cultural competence and cultural humility in the care of diverse patients. *Focus (Am Psychiatr Publ)*. 2020;18(1):49–51. doi:10.1176/appi.focus.20190041
13. Tsuei J, Bandini JI, Thomas AD, et al. A systems-based framework for integrating health equity and patient safety. *Jt Comm J Qual Patient Saf*. 2025;51(7-8):498–506. doi:10.1016/j.jcjq.2025.04.005
14. Chandra S, Mohammadnezhad M, Ward P. Trust and communication in a doctor-patient relationship: a literature review. *J Healthc Commun*. 2018;03(03):36. doi:10.4172/2472-1654.100146
15. American College of Obstetricians and Gynecologists. Effective Patient-Physician Communication. Committee Statement No. 23. Published December 2025. <https://www.acog.org/clinical/clinical-guidance/committee-statement/articles/2025/12/effective-patient-physician-communication>
16. Frank JR, Snell L, Sherbino J. CanMEDS 2015 physician competency framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
17. General Medical Council. Outcomes for graduates. London: General Medical Council; 2018.
18. World Federation for Medical Education. WFME global standards for quality improvement in medical education: basic medical education. London: World Federation for Medical Education; 2015.
19. Kurtz S, Draper J, Silverman J. Teaching and learning communication skills in medicine. 2nd ed. Boca Raton: CRC Press; 2005.
20. Rider EA, Hinrichs MM, Lown BA. A model for communication skills assessment across the undergraduate curriculum. *Med Teach*. 2006;28(5): E127–e134. doi:10.1080/01421590600726540
21. Yu JH, Chang HJ, Kim SS, et al. Effects of high-fidelity simulation education on medical students' anxiety and confidence. *PLoS One*. 2021;16(5): E0251078. doi:10.1371/journal.pone.0251078
22. Veazey K, Notebaert A, Robertson EM. Medical student perceptions of establishing effective clinical communication: a qualitative study. *Adv Health Sci Educ Theory Pract*. 2026;31(2):653–681. doi:10.1007/s10459-025-10468-x
23. Dewi SP, Wilson A, Duvivier R, et al. Perceptions of medical students and their facilitators on clinical communication skills teaching, learning, and assessment. *Front Public Health*. 2023;11:1168332. doi:10.3389/fpubh.2023.1168332
24. Graf J, Loda T, Zipfel S, et al. Communication skills of medical students: survey of self- and external perception in a longitudinally based trend study. *BMC Med Educ*. 2020;20(1):149. doi:10.1186/s12909-020-02049-w
25. Schopper H, Rosenbaum M, Axelson R. "I wish someone watched me interview": Medical student insight into observation and feedback as a method for teaching communication skills during the clinical years. *BMC Med Educ*. 2016;16(1):286. doi:10.1186/s12909-016-0813-z
26. Ullah H, Huma S, Tahir N, et al. Effects of hidden curriculum on medical education and strategies to reshape hidden curriculum as a curriculum promotor: a systematic review. *Adv Med Educ Pract*. 2025;16:2087–2098. doi:10.2147/AMEP.S561248
27. Ammentorp J, Chiswell M, Martin P. Translating knowledge into practice for communication skills training for health care professionals. *Patient Educ Couns*. 2022;105(11):3334–3338. doi:10.1016/j.pec.2022.08.004
28. Bylund CL, Vasquez TS, Peterson EB, et al. Effect of experiential communication skills education on graduate medical education trainees' communication behaviors: a systematic review. *Acad Med*. 2022;97(12):1854–1866. doi:10.1097/ACM.0000000000004883
29. Gilligan C, Powell M, Lynagh MC, et al. Interventions for improving medical students' interpersonal communication in medical consultations. *Cochrane Database Syst Rev*. 2021;(2).
30. Menezes P, Guraya SY, Guraya SS. A systematic review of educational interventions and their impact on empathy and compassion of undergraduate medical students. *Front Med (Lausanne)*. 2021;8:758377. doi:10.3389/fmed.2021.758377
31. Pal AK, Datta D, Venkata Ramulu M, et al. Communication skills in medical education and practice: a bibliometric analysis of the current research landscape. *Cureus*. 2026;18(2): E104018. doi:10.7759/cureus.104018
32. Patel S, Huang M, Miliara S. Understanding treatment adherence in chronic diseases: challenges, consequences, and strategies for improvement. *J Clin Med*. 2025;14(17):6034. doi:10.3390/jcm14176034

33. Tan XH, Foo MA, Lim SLH, et al. Teaching and assessing communication skills in the postgraduate medical setting: a systematic scoping review. *BMC Med Educ.* 2021;21(1):483. doi:10.1186/s12909-021-02892-5
34. Venktaramana V, Loh EKY, Wong CJW, et al. A systematic scoping review of communication skills training in medical schools between 2000 and 2020. *Med Teach.* 2022;44(9):997–1006. doi:10.1080/0142159X.2022.2054693
35. Franco RS, Franco CAGS, Severo M, et al. Reflective writing in the teaching of communication skills for medical students: a systematic review. *Patient Educ Couns.* 2022;105(7):1842–1851. doi:10.1016/j.pec.2022.01.003
36. Janssen L, Schick K, Neurohr T, et al. Reflect to interact: fostering medical students' communication through reflection-focused e-learning. *BMC Med Educ.* 2024;24(1):541. doi:10.1186/s12909-024-05368-4
37. Karnieli-Miller O. Reflective practice in the teaching of communication skills. *Patient Educ Couns.* 2020;103(10):2166–2172. doi:10.1016/j.pec.2020.06.021
38. Lipnevich AA, Kiat KTH. Students' engagement with feedback: current understanding and future directions. In: Lipnevich AA, To J, Kiat KTH, editors. *Unpacking students' engagement with feedback: pedagogy and partnership in practice.* Abingdon: Routledge; 2024. p. 144–155.
39. Savitha D, Anto T, Tv S. Introducing reflective narrative for first-year medical students to promote empathy as an integral part of physiology curriculum. *Adv Physiol Educ.* 2021;45(2):207–216. doi:10.1152/advan.00206.2020
40. Hattie J, Timperley H. The power of feedback. *Rev Educ Res.* 2007;77(1):81–112. doi:10.3102/003465430298487
41. Lipnevich AA, Smith JK. Student–feedback interaction model: revised. *Studies in Educational Evaluation.* 2022;75:101208.
42. Veloski J, Boex JR, Grasberger MJ, et al. Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. *Med Teach.* 2006;28(2):117–128. doi:10.1080/01421590600622665
43. Boud D, Molloy E. Rethinking models of feedback for learning: the challenge of design. *Assess Eval High Educ.* 2013;38(6):698–712. doi:10.1080/02602938.2012.691462
44. Lipnevich AA, Lopera-Oquendo C. Receptivity to instructional feedback: a validation study in the secondary school context in Singapore. *European Journal of Psychological Assessment.* 2022;40(1):22–32.
45. Nicol D, McCallum S. Making internal feedback explicit: exploiting the multiple comparisons that occur during peer review. *Assessment & Evaluation in Higher Education.* 2022;47(3):424–443. doi:10.1080/02602938.2021.1924620
46. Hawe E, Lightfoot U, Dixon H. First-year students working with exemplars: promoting self-efficacy, self-monitoring and self-regulation. *J Furth High Educ.* 2017;43(1):1–15. doi:10.1080/0309877X.2017.1349894
47. To J, Panadero E, Carless D. A systematic review of the educational uses and effects of exemplars. *Assess Eval High Educ.* 2022;47(8):1167–1182. doi:10.1080/02602938.2021.2011134
48. Massey D, Byrne J, Higgins N, et al. Enhancing OSCE preparedness with video exemplars in undergraduate nursing students: a mixed-methods study. *Nurse Educ Today.* 2017;54:56–61. doi:10.1016/j.nedt.2017.02.024
49. Rajlaxmi S, Rao S, Shruthi R, et al. Use of exemplars-based feedforward to improve the performance of medical students in constructed-response assessment. *Adv Med Educ Pract.* 2022;13:123–129.
50. Isaacson JH, Salas RM, Koch C, et al. Measuring medical students' reflection on their learning: modification and validation of the Motivated Strategies for Learning Questionnaire (MSLQ). *BMC Med Educ.* 2018;18(1):248.
51. Marconi K, Muskat M, Catling C, et al. Health professionals' and students' experiences of reflective writing in learning: a qualitative metasynthesis. *BMC Med Educ.* 2021;21(1):407.
52. Tomazin L, Lipnevich AA, Lopera-Oquendo C. Teacher feedback vs. annotated exemplars: examining the effects on middle school students' writing performance. *Studies in Educational Evaluation.* 2023;78:101262. doi:10.1016/j.stueduc.2023.101262
53. Sinkkonen M, Tapani A. Review of the concept “self-regulated learning. *Intern J Social Educ Sci.* 2024;6(1):130–151.
54. Zimmerman BJ. A social cognitive view of self-regulated academic learning. *J Educ Psychol.* 1989;81(3):329–339. doi:10.1037/0022-0663.81.3.329
55. Chi MTH, Wylie R. The ICAP framework: linking cognitive engagement to active learning outcomes. *Educ Psychol.* 2014;49(4):219–243. doi:10.1080/00461520.2014.965823
56. Kember D, McKay J, Sinclair K, et al. A four-category scheme for coding and assessing the level of reflection in written work. *Assess Eval High Educ.* 2008;33(4):369–379. doi:10.1080/02602930701293355
57. Leung YP, Kember D. The relationship between approaches to learning and reflection upon practice. *Educ Psychol (Lond).* 2003;23(1):61–71. doi:10.1080/01443410303221
58. Radović S, Firssova O, Hummel HGK, et al. Improving academic performance: strengthening the relation between theory and practice through prompted reflection. *Active Learning in Higher Education.* 2023;24(2):139–154. Advance online publication. doi:10.1177/14697874211014411
59. Wiggins BL, Eddy SL, Wener-Fligner L, et al. ASPECT: a survey to assess student perspective of engagement in an active-learning classroom. *ar32. CBE Life Sci Educ.* 2017;16(2): Ar32. doi:10.1187/cbe.16-08-0244
60. Lipnevich AA, Mattern K, D'Addario A. Validity evidence for the communication learning assessment: prior coursework as a known group. *Patient Educ Couns.* 2026;149:109602. doi:10.1016/j.pec.2026.109602
61. Mazor KM, King AM, Hoppe RB, et al. Video-based communication assessment: development of an innovative system for assessing clinician–patient communication. *JMIR Med Educ.* 2019;5(1): E10400. doi:10.2196/10400
62. Gale NK, Heath G, Cameron E, et al. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol.* 2013;13(1):117. doi:10.1186/1471-2288-13-117

63. Li S. Measuring cognitive engagement: an overview of measurement instruments and techniques. *IJPES*. 2022;8(3):63–76. doi:[10.52380/ijpes.2021.8.3.239](https://doi.org/10.52380/ijpes.2021.8.3.239)
64. Redmond P, Heffernan A, Abawi L, et al. An online engagement framework for higher education. *OLJ*. 2018;22(1):1175. doi:[10.24059/olj.v22i1.1175](https://doi.org/10.24059/olj.v22i1.1175)
65. Weinstein CE, Mayer R. The teaching of learning strategies. In: Wittrock MC, Handbook of research on teaching. New York: Macmillan; 1986. p. 315–327.
66. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–174. doi:[10.2307/2529310](https://doi.org/10.2307/2529310)
67. Li M, Gao Q, Yu T. Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. *BMC Cancer*. 2023;23(1):799. doi:[10.1186/s12885-023-11325-z](https://doi.org/10.1186/s12885-023-11325-z)
68. Ash SL, Clayton PH. Generating, deepening, and documenting learning: the power of critical reflection in applied learning. *JALHE*. 2009;01(Fall):25–48. doi:[10.57186/jalhe_2009_v1a2p25-48](https://doi.org/10.57186/jalhe_2009_v1a2p25-48)
69. Faza A, Lestari IA. Self-regulated learning in the digital age: a systematic review of strategies, technologies, benefits, and challenges. *International Review of Research in Open and Distributed Learning*. 2025;26(2):23–58.
70. Jasper M. Beginning reflective practice. Cheltenham: Nelson Thornes; 2003.
71. Chaseley TL, Abercrombie S. Using scaffolds to support preservice teachers' reflective practice. *Front Educ*. 2025;10:1621269. doi:[10.3389/feduc.2025.1621269](https://doi.org/10.3389/feduc.2025.1621269)
72. Rolfe G, Freshwater D, Jasper M. Critical reflection for nursing and the helping professions: a user's guide. London: Palgrave Macmillan; 2001.
73. Bjork RA, Bjork EL. Desirable difficulties in theory and practice. *J Appl Res Mem Cogn*. 2020;9(4):475–479. doi:[10.1016/j.jarmac.2020.09.003](https://doi.org/10.1016/j.jarmac.2020.09.003)
74. Hiebert J. The effects of classroom mathematics teaching. *Second Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics*. 2007;1: 371.
75. Mucheli SS, Chow M. When I say... productive struggle. *Med Educ*. 2026;60(9):958–960. doi:[10.1111/medu.70171](https://doi.org/10.1111/medu.70171)
76. Vygotsky LS. Cole M, John-Steiner V, Scribner S, Souberman E. Mind in society: the development of higher psychological processes. Cambridge: Harvard University Press; 1978.
77. Dekkers T, Melles M, Mathijssen NMC, et al. Tailoring the orthopaedic consultation: how perceived patient characteristics influence surgeons' communication. *Patient Educ Couns*. 2018;101(3):428–438. doi:[10.1016/j.pec.2017.08.018](https://doi.org/10.1016/j.pec.2017.08.018)
78. Falkenstein A, Tran B, Ludi D, et al. Characteristics and correlates of word use in physician-patient communication. *Ann Behav Med*. 2016;50(5):664–677. doi:[10.1007/s12160-016-9792-x](https://doi.org/10.1007/s12160-016-9792-x)
79. Miller AN, Duvuuri VNS, Khan M, et al. The relationship of patient ethnicity/race to physician-patient communication: a mixed-methods systematic review. *Health Commun*. 2025;40(11):2165–2190. doi:[10.1080/10410236.2024.2444342](https://doi.org/10.1080/10410236.2024.2444342)
80. Parra Larrotta S, Hernández Rincón EH, Niño Correa D, et al. Effects of the hidden curriculum in medical education: scoping review. *JMIR Med Educ*. 2025;11: E68481. doi:[10.2196/68481](https://doi.org/10.2196/68481)