Joseph Rencic*, Lambert W.T. Schuwirth, Larry D. Gruppen and Steven J. Durning

Clinical reasoning performance assessment: using situated cognition theory as a conceptual framework

https://doi.org/10.1515/dx-2019-0051
Received July 3, 2019; accepted January 7, 2020

Abstract: Developing valid assessment approaches to clinical reasoning performance has been challenging. Situated cognition theory posits that cognition (e.g. clinical reasoning) emerges from interactions between the clinician and situational (contextual) factors and recognizes an opportunity to gain deeper insights into clinical reasoning performance and its assessment through the study of these interactions. The authors apply situated cognition theory to develop a conceptual model to better understand the assessment of clinical reasoning. The model highlights how the interactions between six contextual factors, including assessee, patient, rater, and environment, assessment method, and task, can impact the outcomes of clinical reasoning performance assessment. Exploring the impact of these interactions can provide insights into the nature of clinical reasoning and its assessment. Three significant implications of this model are: (1) credible clinical reasoning performance assessment requires broad sampling of learners by expert raters in diverse workplace-based contexts; (2) contextual factors should be more explicitly defined and explored; and (3) non-linear statistical models are at times necessary to reveal the complex interactions that can impact clinical reasoning performance assessment.

Keywords: assessment; clinical reasoning; situated cognition theory.

***Corresponding author: Joseph Rencic,** MD, Department of Medicine, Boston University School of Medicine, 72 E. Concord Street, Boston, MA 02118, USA, Phone: +1 (617) 638-7468, Fax: +1 (617) 358-7478, E-mail: jrencic@bu.edu
Lambert W.T. Schuwirth: Department of Medical Education, Flinders University, Flinders, Australia
Larry D. Gruppen: Department of Medical Education, University of Michigan, Ann Arbor, MI, USA
Steven J. Durning: Department of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

Introduction

Clinical reasoning has been defined as the cognitive steps leading up to and including establishing the diagnosis and/or treatment of a patient [1]. Although it is a critical component of competence, clinical reasoning assessment has been called the "Holy Grail" of medical education [2], because it is difficult to assess. This difficulty stems from four challenges. First, clinical reasoning cannot be observed directly so it must be inferred from observable behaviors, such as problem solving and decision making [3]. For this reason, we will henceforth use the term, "clinical reasoning performance", to clarify this distinction when referring to its assessment. Second, it is a multi-faceted, complex construct, which means that a variety of assessment tools are likely necessary to obtain a full picture of clinical reasoning performance [4]. Third, it is highly context-specific, as demonstrated by studies that show clinicians' clinical reasoning performances correlate poorly across different occasions even when the patient's clinical findings and diagnosis remain the same [5, 6]. To reliably determine clinical reasoning performance, numerous observations of a variety of chief clinical problems in diverse contexts are necessary. Thus, and fourth, context specificity leads to a feasibility vs. validity trade-off problem, because it is difficult to obtain the numerous direct observations of real clinical encounters, which are felt to be an essential component of valid clinical reasoning performance assessment [7]. More feasible alternatives, such as multiple-choice questions (MCQs), demonstrate some validity evidence, but account for a limited amount of the variance in clinical reasoning performance suggesting a necessary but insufficient role for standardized MCQ examinations in clinician certification [8]. Of these challenges, both the construct of clinical reasoning and context-specificity have been the subject of renewed exploration in the medical education literature. We will review some of the recent literature regarding both in subsequent sections. Although the context-specific nature of clinical reasoning performance has been recognized for decades, researchers have only recently applied theoretical perspectives gleaned from

Published online February 06, 2020

social constructivist theory to gain deeper insights into its mechanisms. Durning et al. [9, 10] created a model that describes clinical reasoning as an emergent phenomenon arising from the interactions of clinicians, patients, and environmental factors. This model has provided a useful framework for exploring how and why context specificity occurs [11]. We believe that this model can be extended to provide a useful theoretical framework for clinical reasoning performance assessment. This paper consists of four sections that provide a justification for this belief: (1) our working definition of the construct of clinical reasoning which highlights its complexity, (2) a historical perspective on the theoretical lenses that have shaped modern views of clinical reasoning and the assessment of its performance, (3) an explanation of why situated cognition theory provides a valuable conceptual framework for clinical reasoning performance assessment, and (4) a description of a situated cognition-based model of clinical reasoning performance assessment with a discussion of its potential utility.

The construct of clinical reasoning

The construct of clinical reasoning is multi-faceted and complex. A recent review of the health professions education literature demonstrated this fact by identifying 110 different words used for clinical reasoning [12]. Constructs of clinical reasoning vary across as well as within health professions because of variations in scope of practice (ME Young, oral communication). Thus, performance standards for clinical reasoning assessment are often health profession- and specialty-specific. For example, unlike a physician, a registered nurse who misdiagnoses but appropriately triages a septic patient is deemed competent because accurate diagnosis is not typically considered within the scope of nursing practice. Likewise, within a given health profession, the diversity of scope of practice creates unique perspectives on the construct of clinical reasoning and the standards of clinical reasoning performance assessment. Table 1 illustrates a comprehensive list of clinical reasoning "tasks" [13, 14], yet a practicing radiologist might only need to demonstrate good performance in 20 of the 24 to be deemed competent in "radiological" clinical reasoning.

Furthermore, clinical reasoning is a complex ability, requiring both declarative and procedural knowledge, such as physical examination and communication skills. Even when clinicians have adequate medical knowledge to solve the case, they may misdiagnose a patient on the basis of poor physical examination skills, for example. Assessment methods vary significantly in the aspects of clinical reasoning that they evaluate and thus very different perspectives on performance can emerge. Consider the non-native English speaking clinician with outstanding medical knowledge who aces standardized multiple-choice examinations but performs poorly on an objective structured clinical examination (OSCE) due to spoken language difficulties. Such an example highlights the importance of assessing all facets of clinical reasoning performance to allow for valid interpretations of competence.

The problem of context specificity

In the 1960s and 1970s, many researchers believed that clinical reasoning performance arose from a clinician's general problem-solving skills (i.e. the use of the scientific method). However, Elstein et al. [15] discovered that experts' processes were no different than novices and their problem-solving processes, though more accurate than novices, varied dramatically across cases (correlations 0.1–0.3). In other words, a maximum of 30% of clinical reasoning performance correlated with a physician's clinical reasoning ability or "traits". They interpreted these data as proof of "content specificity" (i.e. intra-physician differences in content knowledge across different domains). However, a seminal study questioned the notion of content specificity (i.e. medical knowledge) as the sole explanation for clinical reasoning performance variation. Norman et al. [5] demonstrated only moderate correlations in a clinician's performances on *identical* case presentations across two different occasions. This finding and a subsequent study [11] suggested that context influenced a physician's ability to make an accurate diagnosis to an even greater degree than content (i.e. clinician knowledge). Thus, these and other authors argued that problem-solving ability, including clinical reasoning, is largely context-dependent, or context-specific [5, 6, 11, 16].

At that time, context specificity was considered construct-irrelevant variance to be excluded from analysis because it was viewed as "noise" that obscured the true signal of a clinician's clinical reasoning ability. Statistically-speaking, this measurement "error" was quantifiable through generalizability (G) theory [17], which is a commonly used reliability theory that determines the relative contributions of variables of interest, or "facets", to an observed score for a given latent variable.

Table 1: Tasks of clinical reasoning.

Framing the encounter
1. Identify active issues
2. Assess priorities (based on issues identified, urgency, stability, patient preference, referral question, etc.)
3. Reprioritize based on assessment (patient perspective, unexpected findings, etc.)
a. *Consider the impact of prior therapies*
Diagnosis
4. Consider alternative diagnoses and underlying cause(s)
a. *Restructure and reprioritize the differential diagnosis*
5. Identify precipitants or triggers to the current problem(s)
6. Select diagnostic investigations
7. Determine most likely diagnosis with underlying cause(s)
8. Identify modifiable risk factors
a. *Identify nonmodifiable risk factors*
9. Identify complications associated with the diagnosis, diagnostic investigations, or treatment
10. Assess rate of progression and estimate prognosis
11. Explore physical and psychosocial consequences of the current medical conditions or treatment
Management
12. Establish goals of care (treating symptoms, improving function, altering prognosis or cure; taking into account patient preferences, perspectives, and understanding)
13. Explore the interplay between psychosocial context and management
14. Consider the impact of comorbid illnesses on management
15. Consider the consequences of management on comorbid illnesses
16. Weigh alternative treatment options (taking into account patient preferences)
17. Consider the implications of available resources (office, hospital, community, and interand intraprofessionals on diagnostic or management choices
18. Establish management plans (taking into account goals of care, clinical guidelines/evidence, symptoms, underlying cause, complications, and community spread)
19. Select education and counseling approach for patient and family (taking into account patientsand their families levels of understanding)
20. Explore collaborative roles for patient and family
21. Determine followup and consultation strategies (taking into account urgency, how pending investigations/results will be handled)
22. Determine what to document and who should receive the documentation
Selfreflection
23. Identify knowledge gaps and establish personal learning plan
24. Consider cognitive and personal biases that may influence reasoning

From: Goldszmidt et al. [13]. Italicized tasks from McBee et al. [14].

G theory reveals that the observed score in a clinicians (i.e. person, or p facet) performance relates not only to case (*c*) and occasion (*o*) facets but also to the *interactions* (x) between them (e.g. *p* x *c* x *o*) [18]. The person (*p*) component (i.e. the true signal of a clinicians clinical reasoning ability represents) accounted for a relatively small amount of performance variance. The *p* x *c* interaction represents the case specificity facet and the *p* x *c* x *o* interaction the context specificity facet.

From a psychometric perspective, the large unexplained variance in clinical reasoning performance may relate to at least three possibilities: (1) unmeasured factors, or (2) a more complex construct than presumed or (3) truly random statistical variance [19]. Statistics are agnostic about the possibility of unmeasured factors and construct complexity. Thus, a conceptual framework that could provide hypotheses for exploring these possibilities was essential. Dual process theory [20], a dominant clinical reasoning conceptual framework which focuses on information processing within the brain, recognized the influence of some contextual factors, such psychological stress, on thinking but has not traditionally accounted for the environments impact on cognition. However, it was not a natural candidate for exploring context specificity because it was less wellsuited to explore the multifaceted nature of context. Rather, situated cognition theory, which emphasizes that thinking emerges from the complex interactions of multiple contextual factors, seemed to provide greater potential for reformulating conceptions of context specificity, addressing the multifaceted nature of the construct of clinical reasoning, and moving the field forward [21].

Alternative perspectives on context specificity in clinical reasoning: situated cognition theory

Situated cognition theory states that the outcome of cognition, in this case clinical reasoning, is based upon the specifics of the situation, such as physician, patient, and environmental factors [10]. Situated cognition theory posits that the clinician, patient, and environment are interdependent and that clinical reasoning performance emerges from the dynamic interplay of these factors and their interactions [9]. Thus, clinical reasoning performance is not a stable trait inherent to a clinician or a learner, but rather a context-dependent state. This theory aligns well with the empirical data in diagnosis demonstrating context-specific clinical reasoning performance variation [5, 6, 11]. This interdependence of the physician and other contextual factors can provide a possible explanation for why context specificity is seen because although the physician, case (patient), and environment might seem the "same," the occasion is different. Variations in clinician factors (e.g. less stressed), patient factors (e.g. slightly less affable behavior), or environmental factors (e.g. clinical setting, time pressure) will naturally occur on different occasions. Even minor differences in any of these factors have the potential to lead to large changes in the interactions between these components and hence, in performance. In other words, clinical reasoning emerges from these complex interactions, rather than existing solely within a clinician's head. Understanding the change in perspective that situated cognition theory poses on the construct of clinical reasoning, we now turn our attention to its potential impact on clinical reasoning performance assessment.

Situated cognition theory perspective on the assessment of clinical reasoning performance

Durning et al. [11] and Kogan et al. [22] demonstrated the value of using a situated cognition perspective for exploring clinical reasoning performance variation and assessment of resident clinical and interpersonal skills performance through direct observation, respectively. Thus, there are precedents that demonstrate the value of a situated cognition perspective in both clinical reasoning performance and general competence assessment. Situated cognition theory's insights into the context-specificity of diagnosis extend naturally into clinical reasoning performance assessment. The situated cognition theory view recognizes the limitations of the historical assessment approach that views context as "noise" in clinical reasoning performance assessment. A situated cognition theory perspective posits that clinical reasoning performance assessment, like clinical reasoning itself, emerges from a complex interplay between multiple contextual factors and the underlying factors serve as avenues to better understand why context specificity occurs and how clinical reasoning assessment "works" within a given situation. Assessment method, task, and rater factors enter the complex mix of interactions between the clinician (henceforth termed "assessee"), patient, and environment. We propose a model of the interactions between these six components inherent in clinical reasoning performance assessment (Figure 1). We call these components contextual factors because they provide the context for clinical reasoning performance assessment. Because our focus is conceptual, we will not discuss other contextual factors, such as health care team dynamics, the presence of a patient's family members, or institutional culture, which greatly expand and complicate the model.

Our framework further develops previous frameworks in useful ways. The most significant addition to a situated cognition model of clinical reasoning (i.e. clinician, patient, and environment) is the rater. With the exception of tests graded by computer software, a rater significantly impacts clinical reasoning performance assessment. The framework also expands upon Kogan et al.'s [22] situated cognition-related conceptual model of direct observation assessments which was developed by using qualitative analyses of raters' perspectives. Because they limited their study to the mini-CEX direct observation tool, Kogan et al. did not include the assessment task or method in their model. These two contextual factors are more likely to have a significant impact on clinical reasoning performance assessment than other factors that they included in their model (e.g. clinical system, institutional/educational culture). Regardless, these factors would be included in our model within environmental factors. Finally, it highlights the possibility for complex interactions, particularly between raters, assessees, and patients, and the potential value of non-linear statistical analysis in providing deeper insights into context specificity [23].

Research and medical education applications

The 6-faceted model can be useful in both research and medical education. From a research perspective, the

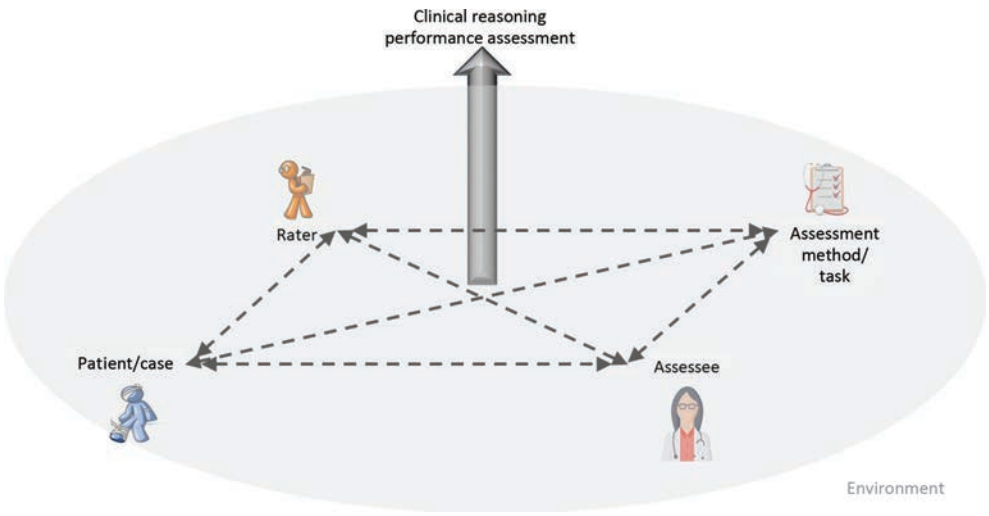

Figure 1: A situated cognition model of clinical reasoning performance assessment.
Clinical reasoning performance assessment emerges from the interactions of various contextual factors, including the assessee, the patient, the assessment method, and the task within a complex physical and cultural environment (gray circle).

model provides clear targets and hypotheses for research designs which might allow for a better understanding of how and to what degree contextual factors impact clinical reasoning performance assessment. One application of the framework is to extend Durning et al.'s prior work [11] with clinical reasoning performance to explore how varying contextual factors, including assessment methods and tasks, affects both the assessee and the rater. Alternatively, assessee performance can be controlled for using scripted clinical encounters to explore how patient and environmental factors, as well as assessment method and tasks, impact rater scoring.

For medical educators, the framework provides a set of contextual factors that can be documented in all assessments. Although some direct observation forms do request documentation of contextual factors (e.g. Mini-CEX), in-training evaluation reports (e.g. end of rotation evaluations) rarely, if ever, require documentation of contextual factors, like average service size, patient acuity or complexity, and environmental factors. Imagine a fellow that is sleep-deprived due to new baby at home on a service with an average census of 18 patients, 30% of which is non-English speaking. A clinical supervisor evaluating this resident could provide a summative rotation narrative and score without mentioning any of these factors that could dramatically impact clinical reasoning performance. Providing contextual questions on assessment forms based on the framework could allow for more nuanced interpretations of clinical reasoning performance assessment by competency committees in residency programs and promotions committees in medical schools. Without such information, raters may fail to observe and comment on key aspects of an assessee's circumstances, leading to poorly-informed assessments of clinical reasoning performance, and misdiagnosis of the cause of a learner's struggles. Thus, such reporting has the potential to improve both summative and formative assessment. One could think of it as the demographics, or (Table 1), of the key components of the clinical reasoning performance (Table 2).

Implications of situated cognition theory on clinical reasoning performance assessment

A situated cognition view has implications on clinical reasoning performance assessment. From a situated cognition perspective, clinical reasoning performance and assessment emerges from the interactions of individual cognition with physical, social, and cultural contexts. This perspective leads to a new and more inclusive definition of clinical reasoning:

> *The phenomena that emerge through the interplay between the cognitive and physical processes of the healthcare professional consciously and unconsciously adapting to interactions with patients and their relations, team members, raters (when applicable), environments, and tasks with the purpose to solve problems and make decisions by continually collecting and interpreting patient data, prognosticating, weighing the benefits and risks of actions, and understanding patient preferences in order to develop a diagnostic and therapeutic management plan that aims to improve a patient's well-being.*

Table 2: Example of elements of an in-training end-of-rotation evaluation form for clinical reasoning performance from situated cognition perspective.

Assessee factors: ☐ Rotation experience ☐ Average census ☐ Max census (# of days) ☐ Assessee personal problems during rotation ☐ Other	Patient factors: ☐ Average patient acuity (low, mod, high) ☐ Average patient complexity (low, mod, high): ☐ Other	Environment factors: ☐ Overnight call ☐ Typical time duration for new patient evaluation ☐ Resident/intern supervision quality	Your experience: ☐ Months per year working with students on average ☐ Years as teaching attending ☐ Total hours of student observation ☐ # of chief complaints observed ☐ # of times clinical reasoning tasks observed: History___, Physical examination___, Problem representation___, Prioritization of differential diagnosis___, Justification of diagnosis___ ☐ Other___

Make informed diagnostic and therapeutic decisions that result in optimal clinical judgment[a]	Recalls and presents clinical facts in the history and physical in the order they were elicited without filtering, reorganization, or synthesis; demonstrates analytic reasoning through basic pathophysiology results in a list of all diagnoses considered rather than the development of working diagnostic considerations, making it difficult to develop a therapeutic plan	Focuses on features of the clinical presentation, making a unifying diagnosis elusive and leading to a continual search for new diagnostic possibilities; largely uses analytic reasoning through basic pathophysiology in diagnostic and therapeutic reasoning; often reorganizes clinical facts in the history and physical examination to help decide on clarifying tests to order rather than to develop and prioritize a differential diagnosis, often resulting in a myriad of tests and therapies and unclear management plans, since there is no unifying diagnosis	Abstracts and reorganizes elicited clinical findings in memory, using semantic qualifiers [such as paired opposites that are used to describe clinical information (e.g. acute and chronic)] to compare and contrast the diagnoses being considered when presenting or discussing a case; shows the emergence of pattern recognition in diagnostic and therapeutic reasoning that often results in a well synthesized and organized assessment of the focused differential diagnosis and management plan	Reorganizes and stores clinical information (illness and instance scripts) that lead to early directed diagnostic hypothesis testing with subsequent history, physical examination, and tests used to confirm this initial schema; demonstrates well-established pattern recognition that leads to the ability to identify discriminating features between similar patients and to avoid premature closure; Selects therapies that are focused and based on a unifying diagnosis, resulting in an effective and efficient diagnostic work-up and management plan tailored to address

[a]From: Pediatrics Milestone Project Working Group. The Pediatrics Milestone Project. Available at: https://www.acgme.org/Portals/0/PDFs/Milestones/PediatricsMilestones.pdf. Accessed December 11, 2019.

This definition adds some of the contextual factors excluded from the model for simplicity to show how complex the assessment of clinical reasoning performance can be. In conformity with Young et al.'s findings [12], this phrase should not be viewed as a "one size fits all" definition, but one which can be used by researchers and educators using a situated cognition lens to clarify their perspective. A situated cognition perspective leads to a different assessment focus as compared to the more traditional information processing view (e.g. dual process theory) (Table 3).

From a situated cognition perspective, workplace-based assessments have the potential to provide the most authentic evidence regarding clinical reasoning

Table 3: Comparison of information processing versus situated cognition perspectives on assessment.

Theory	How assessment occurs	Factors influencing assessment	What are implications for construction of assessment?
Information processing	Content focus, sociocultural context irrelevant	Learner and rater mental activities	• Emphasize knowledge content and organization
Situated cognition	Situational focus, performance emerges within sociocultural context	Interactions between assessee, patient, and rater with each other, as well as the environment, assessment method, and task	• Recognize assessment as a sociocultural phenomenon • Promote assessment in authentic situations across varied clinical contexts

performance competence because of the richness of the interactional factors at play. Only workplace-based assessments provide information on authentic clinical performance within the complex context of assessee, rater, patient, and environmental interactions. This recognition endorses direct observation of patient encounters as a fundamental method of assessing clinical reasoning performance, and parallels similar endorsements from the competency-based medical education community [7]. The emphasis on direct observation highlights the need for expert raters. Without significant training, such raters will provide limited assessments at best and inaccurate ones at worst. One caveat to this view is the notion that credible assessment requires developmentally-appropriate contexts. Assessing a second-year medical student's ability to stabilize a critically-ill patient is not warranted. Situated cognition provides a clear theoretical framework for medical educators to understand the conceptual basis for these recommendations.

Embracing complexity: a model for clinical reasoning assessment

Historically, psychometric theory, specifically G theory, provided critical insights into the clinical reasoning assessment literature by illuminating the importance of the contribution of different contextual factors and their interactions. Psychometric theory continues and will continue to be an essential component of clinical reasoning assessment. However, it viewed these factors as predefined, stable, distinguishable components of clinical reasoning assessment, describing the percentage of performance variation attributable to each facet and its interactions. In contrast, situated cognition theory posits that these factors are interdependent, dynamic (i.e. not predefined), and interact with each other in non-linear, and often complex, ways [23]. They cannot be measured as stable or fixed "facets" within a linear model such as G theory. Exploring the complex interactions of these contextual factors through non-linear analysis may enhance understanding of clinical reasoning performance variation and its assessment.

Future directions

Our model of clinical reasoning performance assessment (Figure 1) provides a framework for future investigators to explore pentadic, hexadic, and even more complex interactions between the components of clinical reasoning assessment. This research may require nonlinear approaches as discussed because many of these interactions are complex and not predictable by linear models [23]. Continued study and development of clinical reasoning assessment methods and tools is essential. As previously mentioned, current clinical reasoning assessment methods should be modified to incorporate in-depth descriptions of contextual factors. Novel tools, such as self-regulated learning with microanalytic techniques and biological assessment tools like functional magnetic resonance imaging (MRI), require further study to determine their role within the clinical reasoning performance assessment armamentarium and to enhance their scalability. Furthermore, as virtual reality gains a larger footprint in medical education, there will be the opportunity to vary the different components of clinical reasoning performance assessment, while measuring biological parameters of assessees and raters to assess factors such as their stress and cognitive load. The dearth of therapeutic reasoning assessment methods also must be remedied [24]. The therapeutic clinical reasoning literature resides primarily within the medical decision-making community with a focus on mathematical modeling using Bayesian reasoning to create formal decision analyses. With its recognition of multiple pathways and possible treatments for a given clinical problem, situated cognition

theory can provide a perspective that will promote a multidisciplinary approach to explore therapeutic reasoning and its assessment.

Thinking beyond assessment of learning, another major benefit of exploring the mechanisms of context specificity and clinical reasoning performance variation is assessment *for* learning [25]. Understanding clinical reasoning performance variation may allow educators to choose assessment methods that maximize learning as well as validity. Future studies should incorporate learning outcomes as a specific measure of the utility of a clinical reasoning performance assessment methods.

Conclusions

The concept of clinical reasoning has changed dramatically over the last 40 years from a process-dependent trait to a dynamic state affected by complex, non-linear interactions between the clinician, patient, and the environment (i.e. situated clinical reasoning). Clinical reasoning performance assessment exacerbates this complexity by adding an assessment method, task, and rater into the mix of clinician, patient, and environment. The current approach to clinical reasoning performance assessment often focuses on single right diagnoses or treatments, ignoring the richness and diversity of process and outcomes which can provide a deeper understanding of its variation. By exploring clinical reasoning performance assessment through the lens of situated cognition theory new insights emerge. We believe that our model may promote research that explores the mechanisms of context specificity in the assessment of clinical reasoning performance and modified assessment methods that improve the credibility evidence with which we make high-stakes assessment decisions regarding clinical reasoning performance.

Acknowledgments: We would like to acknowledge the outstanding team of medical educators who participated in this literature review: Tiffany Ballard, Michelle Daniel, Carlos Estrada, David Gordon, Anita Hart, Eric Holmboe, Valerie Lang, Katherine Picho, Temple Ratcliffe, Sally Santen, Ana Silva, and Dario Torre. We also would like to acknowledge our librarians, Nancy Allee, Donna Berryman, and Elizabeth Richardson, who aided in developing our search strategy.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Holmboe ES, Durning SJ. Assessing clinical reasoning: moving from in vitro to in vivo. Diagnosis 2014;1:111–7.
2. Schuwirth L. Is assessment of clinical reasoning still the Holy Grail? Med Educ 2009;43:298–300.
3. Kreiter CD, Bergus G. The validity of performance-based measures of clinical reasoning and alternative approaches. Med Educ 2009;43:320–5.
4. Young M, Thomas A, Lubarsky S, Ballard T, Gordon D, Gruppen LD, et al. Drawing boundaries: the difficulty in defining clinical reasoning. Acad Med 2018;93:990–5.
5. Norman GR, Tugwell P, Feightner JW, Muzzin LJ, Jacoby LL. Knowledge and clinical problem-solving. Med Educ 1985;19:344–56.
6. Eva KW. On the generality of specificity. Med Educ 2003;37:587–8.
7. Holmboe ES. Realizing the promise of competency-based medical education. Acad Med 2015;90:411–3.
8. Lipner RS, Hess BJ, Phillips RL. Specialty board certification in the United States: issues and evidence. J Contin Educ Health Prof 2013;33(Supp 1):S20–35.
9. Durning SJ, Artino Jr AR, Pangaro LN, van der Vleuten C, Schuwirth L. Perspective: redefining context in the clinical encounter: implications for research and training in medical education. Acad Med 2010;85:894–901.
10. Durning SJ, Artino AR. Situativity theory: a perspective on how participants and the environment can interact: AMEE guide no. 52. Med Teach 2011;33:188–99.
11. Durning SJ, Artino AR, Boulet JR, Dorrance K, van der Vleuten CP, Schuwirth LW. The impact of selected contextual factors on internal medicine experts' diagnostic and therapeutic reasoning: does context impact clinical reasoning performance in experts. Adv Health Sci Educ Theory Pract 2012;17:65–79.
12. Young ME, Thomas A, Gordon D, Gruppen LD, Lubarsky S, Rencic J, et al. The terminology of clinical reasoning in health professions education: implications and considerations. Med Teach 2019;41:1277–84.
13. Goldszmidt M, Minda JP, Bordage G. Developing a unified list of physicians' reasoning tasks during clinical encounters. Acad Med 2013;88:390–7.
14. McBee E, Ratcliffe T, Goldszmidt M, Schuwirth L, Picho K, Artino AR Jr, et al. Clinical reasoning tasks and resident physicians: What do they reason about? Acad Med 2016;91:1022–8.
15. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving: an analysis of clinical reasoning. Cambridge, MA: Harvard University Press, 1978.
16. Perkins D, Salomon G. Are cognitive skills context-bound? Educ Res 1989;18:16–25.
17. Shavelson RJ, Webb NM, Rowley GL. Generalizability theory. Am Psychol 1989;44:922.
18. Shavelson RJ, Baxter GP, Gao X. Sampling variability of performance assessments. J Educ Meas 1993;30:215–32.

19. Schauber SK, Hecht M, Nouns ZM. Why assessment in medical education needs a solid foundation in modern test theory. Adv Health Sci Educ 2018;23:217–32.
20. Kahneman D. Thinking, fast and slow. New York: Farrar, Straus and Giroux, 2011.
21. Durning SJ, Artino Jr AR, Schuwirth L, van der Vleuten C. Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. Acad Med 2013;88(4):442–8.
22. Kogan JR, Conforti L, Bernabeo E, Iobst W, Holmboe E. Opening the black box of clinical skills assessment via observation: a conceptual model. Med Educ 2011;45:1048–60.
23. Durning SJ, Lubarsky S, Torre D, Dory V, Holmboe E. Considering "nonlinearity" across the continuum in medical education assessment: supporting theory, practice, and future research directions. J Contin Educ Health Prof 2015;35(3): 232–43
24. Cook DA, Sherbino J, Durning SJ. Management reasoning: beyond the diagnosis. J Am Med Assoc 2018;319: 2267–8.
25. van der Vleuten CP, Schuwirth L, Driessen E. A model for pro- grammatic assessment fit for purpose. Med Teach 2012;4: 205–14.