Teaching and Learning in Medicine, 25(S1), S26–S32

ISSN: 1040-1334 print / 1532-8015 online
DOI: 10.1080/10401334.2013.842911

Diagnostic Reasoning: Where We've Been, Where We're Going

Sandra M. Monteiro
Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada

Geoffrey Norman
Department of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, Ontario, Canada

Recently, clinical diagnostic reasoning has been characterized by "dual processing" models, which postulate a fast, unconscious (System 1) component and a slow, logical, analytical (System 2) component. However, there are a number of variants of this basic model, which may lead to conflicting claims. This paper critically reviews current theories and evidence about the nature of clinical diagnostic reasoning. We begin by briefly discussing the history of research in clinical reasoning. We then focus more specifically on the evidence to support dual-processing models. We conclude by identifying knowledge gaps about clinical reasoning and provide suggestions for future research. In contrast to work on analytical and nonanalytical knowledge as a basis for reasoning, these theories focus on the thinking process, not the nature of the knowledge retrieved. Ironically, this appears to be a revival of an outdated concept. Rather than defining diagnostic performance by problem-solving skills, it is now being defined by processing strategy. The version of dual processing that has received most attention in the literature in medical diagnosis might be labeled a "default/interventionist" model,[17] which suggests that a default system of cognitive processes (System 1) is responsible for cognitive biases that lead to diagnostic errors and that System 2 intervenes to correct these errors. Consequently, from this model, the best strategy for reducing errors is to make students aware of the biases and to encourage them to rely more on System 2. However, an accumulation of evidence suggests that (a) strategies directed at increasing analytical (System 2) processing, by slowing down, reducing distractions, paying conscious attention, and (b) strategies directed at making students aware of the effect of cognitive biases, have no impact on error rates. Conversely, strategies based on increasing application of relevant knowledge appear to have some success and are consistent with basic research on concept formation.

Keywords clinical reasoning diagnosis, dual processing, cognition, problem-solving

Correspondence may be sent to Geoffrey Norman, McMaster University, Program for Educational Research and Development, Room 3519, MDCL, 1200 Main Street West, Hamilton, ON L8N 3Z5, Canada. E-mail: norman@mcmaster.ca

INTRODUCTION

Perhaps the earliest model of clinical reasoning was the hypothetico-deductive theory,[1] which equated diagnostic expertise to a specialized thinking process or ability to quickly generate a diagnostic hypothesis early in the clinical encounter. However, this early model was abandoned in the 1980s, as detailed analyses revealed that both experts and novices generated early hypotheses largely as a strategy to reduce cognitive load, a consequence of the limited size of working memory. Moreover there was very little evidence that the process was central to diagnostic expertise.[1,2] Experts did not generate more hypotheses earlier; they just generated better hypotheses. In short, expertise resided in content knowledge, not process.

Patel and Groen[3] then advanced a theory derived from the artificial intelligence literature that experts used "forward-reasoning" from data to diagnoses, and novices were more likely to use "backward-reasoning" from hypotheses to data. However, others[4] have shown that the link may be an artifact of the experimental situation more than a marker of expertise. Moreover, forward and backward reasoning strategies do not lend themselves to the development of strategies for medical education or diagnostic error reduction; it is not likely that admonishing students to improve their forward reasoning skills will transform them into experts.

Influenced by research in psychology and sociology, the field shifted and later theories postulated various forms of semantically rich medical knowledge—semantic axes,[5] illness scripts,[6,7] probability matrices,[8] propositional networks.[9] In contrast to the theories that described the process of hypothesis generation or forward reasoning, these theories identified knowledge representations as sources of hypotheses and central to expertise.

More recently, there has been renewed interest in the process (as opposed to the content) of diagnostic reasoning, especially in the context of diagnostic errors. Currently, the dominant model is a dual process framework,[10] in which reasoning may proceed by a fast, unconscious, retrieval process (System 1) or a more analytical, slow, deliberate, and conscious logical process (System 2). Increased reliance on System 2 processing is assumed to improve diagnostic reasoning; the quality of knowledge representations is rarely questioned, as the focus has been less on modeling diagnostic expertise, then on identifying cognitive biases associated with diagnostic errors in reasoning.[10–12]

In the present article, we contrast this perspective with the view, derived from a different tradition in psychology, that medical diagnosis is a categorization and memory task[13,14] dependent on analytical and experiential knowledge. From this "memory" perspective, diagnoses are not reasoned so much as they are recognized. In other words, diagnosis is primarily recognition of similar and familiar previously seen signs and symptoms. The processes involved are complex, but considered secondary, and only serve to describe how memory functions in accessing prior knowledge. Learning based on improving the quality of knowledge representations is fundamental to this perspective, as it facilitates the recognition of symptoms and disease categories.

By outlining how these two perspectives have developed, we hope to provide a better understanding of what they have to offer. We begin by reviewing the influence of reasoning theories in medical education and the evidence for strategies to improve diagnostic reasoning. We then review the history of research based on memory theories and the evidence for strategies to improve diagnostic memory and learning. Finally, we conclude with a proposal to combine the two programs to improve learning of reasoning in medical education.

THEORIES OF REASONING STRATEGIES IN PSYCHOLOGY

Traditionally, research in problem solving and decision making was concerned with general thinking processes. Researchers were less interested in the contribution of memory and experience and more interested in how a solution was derived logically from the data.[15] To reduce effects of individuals' prior experiences, these thinking processes were frequently studied using artificially created probability problems (e.g., the famous Linda problem of Tversky & Kahneman).[16] Paradoxically, although the goal was to understand how humans solve problems in the absence of specialized knowledge, most people found these contrived problems difficult to solve without specialized and extensive training in logic and probability—a different kind of specialized knowledge.[17,18]

These studies were interpreted as demonstrating that people did not rely on normative logic but instead recruited simple rules or "heuristics," which are prone to various "cognitive biases."[17] Human reasoning is often described as irrational or subrational because of reliance on heuristics and biases, which result in suboptimal or incorrect responses, *particularly for the artificial problems specifically designed to result in a failure of heuristics*. Despite its tautological nature, this "heuristics and biases" research program has a strong following in medical education.[10–12,19,20]

Dual Process Framework

The research on heuristics and biases is closely aligned with a dual process model of reasoning briefly defined earlier. The general description of this model, consisting of two independent systems, has become common knowledge with the popularity of Kahneman's best-selling book *Thinking Fast and Slow*.[17] However, this model's specific theoretical framework, called a "Default-interventionist" model, is quite problematic for practical applications.

Default-interventionist models of reasoning propose separate systems that operate in stages (one after the other) or exclusively (one or the other).[17,21,22] In the default-interventionist model, the default mode of reasoning is System 1, which contributes to reasoning by rapidly recruiting heuristics. System 2, on the other hand, intervenes when difficulty or bias arises, yet is underutilized because of cognitive demands.[17,21] As heuristics are associated with biases and are generally viewed as suboptimal strategies, System 1 is similarly viewed as suboptimal and error prone.[10,17,20,23,24] There is an inherent assumption within this framework that the processes are independent and capable of *cognitive decoupling*—the capacity for System 2 to independently assess, manipulate, or even inhibit information retrieved by System 1.[18] This concept of cognitive decoupling is critical to the assumption that control of one or the other system is under conscious influence, leading several authors to conclude that reliance on heuristics can be discouraged through training to consciously recruit System 2 and normative rules, quite independent of content knowledge and experience.[17,25,26]

Interventions to Reduce Error

Within this framework, attempts to devise interventions to reduce errors have focused on three broad strategies that are assumed to increase reliance on System 2 processing—slowing down, reflection, and cognitive forcing.[27]

Slowing down. Kahneman[17] proposed that System 1, although efficient, is inherently flawed, as it contributes to basic pattern recognition and does not engage logical reasoning processes. Because System 1 has been characterized by a faster speed of processing compared to System 2, perhaps the simplest intervention is to simply admonish subjects to go slow, be systematic, be thorough, take their time, and so on. Such an intervention follows directly from the dual process framework, and should result in increased reliance on analytical, System 2, processes. It also seems to be common sense that more time spent reasoning through a problem should result in better (more accurate) solutions.

Surprisingly, this general assumption that faster responses are more error prone than slower responses is not supported by research in diagnostic reasoning. One observational study showed that accuracy was associated with *faster*, not slower, response times.[28] Further experimental studies testing an intervention directed at slowing down did show increased response time but no effect on accuracy.[29,30] Finally a study, in which distractions were introduced to hinder System 2 processing during diagnosis, showed a small increase in response time but no effect on accuracy.[31] In light of the discussion of the previous section, all of these interventions amount to manipulation of the amount of cognitive resources allocated to System 2 thinking, yet all have shown no benefit. The idea, that errors can be eliminated by slowing down to increase reliance on System 2 thinking, finds no support from these studies.

Cognitive forcing strategies. Consistent with heuristics and biases research, Croskerry[11,12] instructed both novice and expert physicians to proceed slowly and be aware of more than 30 sources of cognitive bias. *Cognitive forcing strategies* are a set of warnings that are designed to encourage metacognition (a heightened analytic inspection of one's own thought processes) for the purposes of preventing errors. The promise of thinking strategies that reduce errors is very appealing; however, cognitive forcing strategies have not been shown to be beneficial.[27,32,33]

Cognitive forcing strategies devalue the role of experiential and formal knowledge by discouraging reliance on hypotheses to guide the identification of relevant symptoms; the view is that all errors result from several distinct cognitive biases, which, if eliminated, would reduce errors. Defining these discrete elements of cognition as responsible for error is a one sided argument. As one example, "confirmation bias," the active pursuit of data to support an initial hypothesis, is viewed as a common source of error. However, this process may be an asset. For instance, a working hypothesis affects the relevance of features in a written medical case.[34,35] Brooks et al.[36] demonstrated that feature lists contained more accurate items when participants had a working hypothesis, and Norman et al.[37] showed that overall diagnostic accuracy was improved when participants followed a form of "backward reasoning" by starting with a hypothesis. These studies suggest that the active pursuit of supportive data is a useful process, often facilitated by a correct initial hypothesis, and labelled as "confirmation bias" only in the event that the end result is an error.

Reflective practice. The commonsense assumption that conscious reflection can improve judgment also fits within a default interventionist framework. The concept of reflective thought was popularized in medicine by *The Reflective Practitioner*,[38] which has two forms. Increased awareness and introspection while treating a patient comprise reflecting *in* action, whereas retrospective analyses of decisions comprise reflecting *on* action.

Several studies have examined the influence of reflection on the accuracy of hypothesis generation.[38–41] Mamede et al.[39] manipulated the ambiguity of several cases in a diagnostic task to induce unstructured reflective thought but found that it did not improve accuracy. A structured reflective method was then tested where the participants listed possible diagnoses, identified critical features, then completed a matrix relating features to diagnoses, and eventually arrived at a conclusion. This method, designed to support a deliberate, System 2 approach to diagnosis has had mixed results.[39–42] In one study it reduced errors only for diagnoses where the potential for error was pointed out.[40] In two studies, although the reflective method did not provide an overall advantage, it did improve accuracy for a subset of cases that the authors identified as difficult.[39,40] In all these studies however, any advantage provided by the reflective method was quite small and idiosyncratic to the study, despite the fact that the method was based on a highly structured and intensive procedure.

The authors justify the structure and intensity of their procedure by comparing their method to that of deliberate practice proposed by Ericsson and Charness.[43,44] However, the characteristics of deliberate practice may have different effects on novice compared to expert performers.[45] The benefits of a structured program based on the principles of deliberate practice may be limited to the early stages of learning and may prove too cumbersome to maintain once expertise has been achieved. There is also evidence to suggest that a reflexive process is recruited for areas of expertise, whereas reflection is reserved only for unfamiliar content.[46,47] It is conceivable that forcing the use of reflective processes prevents a physician from recruiting prior experience, from their primary asset, or from feeling like an expert. Moreover, there is evidence that more intuitive diagnoses can be highly accurate,[37,48] calling into question the need for such measures.

One final point about the role of reflection. In contrast to the "cognitive forcing strategy" approach, which views errors as a consequence of general "hardwired" cognitive biases, reflection strategies explicitly mobilize analytical knowledge, encouraging subjects to explicitly consider diagnostic alternatives and consider the relation between diagnoses and case features. Thus, although the intervention may be viewed as a process strategy, equally it relies heavily on the clinician's specific knowledge.

Conceptual problems with biases and the default-interventionist model. In summary, the evidence suggests that strategies to increase awareness of reasoning biases are neither necessary nor sufficient for reducing errors in medicine. Conversely, reflection strategies that mobilize relevant knowledge show some benefits.

There are several tenuous assumptions associated with the model. First, the notion that the two systems are "decoupled" and under conscious control may be incorrect. System 1 processing amounts to retrieval of knowledge from memory, and likely proceeds on a time scale of the order of hundreds of milliseconds. Conscious attempts to "speed up" or "slow down" may well alter the amount of cognitive resources devoted to System 2, analytical reasoning, but are unlikely to influence the rapid retrieval processes of System 1. Second, it is fallacious to associate cognitive bias solely with System 1. Biases like "confirmation bias," actively seeking information to rule in a diagnosis, "anchoring and adjustment"—changing probabilities of outcomes by adjusting from a baseline, "premature closure"—arriving at a conclusion without accounting for critical information, all arise during the process of gathering additional data and explicitly weighting alternatives, which is a conscious, System 2, activity.

More fundamentally, to presume that errors derive solely from defects in strategy, as opposed to defects in knowledge, is inconsistent with two general and robust findings in the literature on diagnostic reasoning. First, cognitive biases are supposed to be general and skill-like(or unskill-like), yet decades of research consistently identify "content-specificity" in problem solving, where performance on one problem is a poor predictor of performance on another.[1] Second, there is very little evidence that cognitive biases can be influenced by training, and some evidence that they cannot (as previously described). Yet clearly, medical education "works"—practicing physicians are better diagnosticians than residents who are, in turn, better than students.[49] Yet none may even be aware of their cognitive biases, and few are likely to have had specific training in debiasing strategies (presuming that such training would be effective).

These fundamentally weak assumptions have led to disagreement among different dual process theorists, and even between different arguments from the same authors.[21,50–53] Although Evans and Stanovich[21] argued that both types of processing could lead to errors and that it would be incorrect to assume that one process was superior, elsewhere[22,53] they argue that System 2 is associated with rationality and intelligence, which have been interpreted to mean that System 2 is superior.

Although many authors distance themselves from these assumptions, it is difficult to disentangle the architecture of a default interventionist dual systems model from the assumption that System 1 is inferior and System 2 is superior. This confusion in the literature has led many to consider abandoning dual process models altogether.[50,53–55]

Not all dual process frameworks are problematic however. Parallel-competitive models propose separate processes that influence each other and operate simultaneously.[22] Parallel operating frameworks are used in models of memory such as categorization[56] and recognition,[57] in which each process is linked to a different form of knowledge: nonanalytic experiential knowledge and analytic rule-based knowledge. The process dissociation method was developed to calculate the relative contribution of these parallel processes, as within this framework neither process is considered superior and no task is considered a pure measure of a single mode of processing.[58] These dual process models, focusing on access of memory, may be better suited for application to medical education and strategies to reduce error.

The Role of Memory in Reasoning

In contrast to research on reasoning, research on memory focuses on how information is encoded and recalled to solve problems and maintain goals. Performance on each new problem or task is assessed in relation to prior experience, which may be a characteristic of individual experience, as in research with chess experts,[59] or may be experimenter controlled, where participants learn materials specific to the experiment (e.g., lists of words).[14] More important, research on human memory identifies several complex factors affecting recall errors (e.g., false memory, failure to recognize, etc.) and rarely emphasizes reasoning-related errors. For example, the phrasing of a question can influence whether an object is recognized as old or new,[60] and contextual information can cue the recall of different information, affecting accuracy.[61,62] Therefore, memory researchers propose learning strategies to ensure that complex factors such as context have a reduced impact on recall accuracy.[62,63]

We return to memory-based strategies to improve learning, but for now we focus on discussing two parallel-processing models of memory that are relevant to improving our understanding of medical diagnosis: categorization and recognition.

Categorization. The ability to identify objects in our environment is explained by models of categorization and concept formation. There are two classes of categorization models: prototype and exemplar.[56] The main distinction is that, in a prototype model, the role of memory is to retrieve an abstracted average representation of individual experiences; in the exemplar model, the role of memory is to retrieve relevant closely matched individual experiences and their features.

Current literature on categorization continues to debate the relevance of prototype versus exemplar models.[64,65] For the purpose of this review, we do not take a side in this debate, as we believe that both forms of knowledge are important for different aspects of learning. For example, novices are possibly limited to reliance on prototypical or analytic knowledge, having few exemplars or experiences to draw upon. As well, both prototypes and exemplars can be shown to influence categorization, largely as a function of experimental design.[66] Regardless of the model, accuracy in categorization is undoubtedly related to the quality and quantity of experiences.

We have already mentioned that medical diagnosis has been described as a categorization task,[67] and there is sufficient evidence supporting the theory that similar prior experiences influence diagnostic accuracy. However, the identification of previously seen disease categories by reliance on similar prior exemplars is sometimes disparagingly referred to as "basic" pattern recognition.[10] We propose that recognition is itself a complex process that contributes to categorization and accurate medical diagnosis.

Recognition. A model of recognition memory proposes that a previously seen item or person can be recognized through a recollection process that recruits specific details about the prior event and a familiarity based process that is more vague.[57,68] For example, recognizing that a pattern of symptoms has been seen before can be accomplished by recalling the details of an identical description in a textbook, by a vague sense that a similar pattern has been seen before on another patient, or by some combination.[57,63] Of importance, both processes are thought to contribute in parallel to recognition memory, although in varying proportions, and there is no known method for measuring the pure influence of any one process as both processes are equally susceptible to environmental or contextual influences, therefore equally susceptible to error.[58,69,70] Learning new information also relies on previously stored experiences, as familiarity can activate previously stored knowledge, whereas recollection can facilitate associations between old and new information in memory and help define how new information is perceived.[57,71]

There is clearly some overlap between a memory-based model of reasoning involving experiential and analytical knowledge forms and dual-processing models, where System 1 amounts to retrieval of prior experiences and System 2 is related to application of rules relating features to categories. However, the emphasis in memory models is not on the process of retrieval but on the nature of the knowledge retrieved.

Memory-Based Strategies to Enhance Reasoning

Based on these principles of memory, there are several well-established strategies that can improve learning. Here we discuss two strategies that have already been investigated and demonstrated to be successful in a medical education context: test-enhanced learning and mixed practice.

Test enhanced learning. The value of testing for assessment is well known, and increasingly so is the value of testing for learning. Testing has been shown to improve retention for material compared to not being tested or studying alone.[71] The testing effect or "test enhanced learning" effect has been demonstrated in several studies by Roediger and colleagues, both in general and medical education.[71] The effect can be understood within the recognition model, whereby retrieval of information from long-term memory (i.e., recollection) can facilitate the learning of new information. More generally, the act of repeatedly retrieving information strengthens associations in memory by providing contextual variability and increasing meaning. Furthermore, tests that require the production of answers (e.g., short answer, fill-in-the-blank, essay) have led to better retention compared to multiple-choice tests,[71] which is consistent with findings in psychology that recollection is improved for items that were generated during learning.[58]

Mixed practice. To the extent that effective reasoning, based on a memory model, is largely derived from an extensive experiential and analytical knowledge base, the emphasis for strategies to improve reasoning skills changes from practicing a process to acquiring examples. Although the importance of practice has been emphasized in the literature on deliberate practice, memory models go further than simply examining the amount of practice. In particular, the categorization task of diagnosis requires learning those features that discriminate one category from another. One strategy is acquisition of examples from "mixed practice," where confusable examples are learned together and features to distinguish one from another are examined. In one study, mixed practice was contrasted to "blocked" practice (one category at a time) in the teaching and practice of ECG interpretation skills. Students who learned by mixed practice showed a 17% increase in test scores compared to students who practiced in the traditional blocked method.[72]

A CHANGE OF PERSPECTIVE IS DUE

Research into cognitive biases contributing to diagnostic errors in medicine has had a disproportionately strong influence on recent research on clinical reasoning, primarily as a consequence of retrospective analyses of diagnostic errors that estimated that 74% of fatal diagnostic errors arose from reasoning biases like premature closure and availability bias.[73,74] Although these retrospective analyses may themselves be susceptible to hindsight bias, these error rates attract a great deal of attention.[75] A consequence is that a significant portion of medical education literature has been dedicated to identifying the sources of cognitive errors, rather than identifying the best strategies for learning the prerequisite knowledge to avoid errors. As Croskerry[76] stated,

> Most of our thinking in clinical practice is of the inductive type, and we should understand its nature and limitations. Inductive thinking is the logic of experience. . . . Those who have learned well from experience are said to have clinical acumen, and while there is no substitute for experience, we might shorten the road by teaching some of the basic flaws and biases known to be present in everyday thinking. (p. 1226)

He described the nature of expertise and correctly acknowledged the importance of experience in reducing errors. However he drew the conclusion that attention should be focused on teaching "flaws and biases." Errors are viewed only as unnecessary and preventable problems that can be eliminated by improving reasoning skills.[10–12] To this end, diagnostic errors are typically associated with cognitive biases in reasoning rather than gaps in knowledge.[10–12,19,20,76]

However, this attribution to cognitive biases is only one possibility. A recent retrospective study[77] of chart reviews similar to the Graber[78] study using a different taxonomy did not identify thinking errors derived from cognitive biases. Instead, they found that most errors (58%) were "mistakes" defined as "an intended act, but the physician does not *know* [emphasis added] it is incorrect" (p. 151). Errors were related to lack of knowledge, not bias.

It is through the process of applying knowledge, making mistakes, and learning from them that novices become experts. Errors are a necessary element of early learning, focusing learning on the weaknesses in knowledge identified by errors ensures that future opportunities for error are reduced during professional practice.[79] The benefits of learning from errors can be gained through the strategic use of testing and mixed practice.

We argue that it is insufficient to target the cognitive biases that are purported to act as sources of diagnostic error while ignoring the value of prior experience and knowledge. If physicians spend more time practicing getting it right, and acquiring experiential knowledge during training, they should make fewer errors later on. Strategies based on explication of cognitive biases have, to date, have been shown to be ineffective.[31,32,80] A new approach is needed; we suggest incorporating strategies grounded in memory research.

To accomplish this, we cannot rely on a natural distribution of medical problems to expose physicians to the variety of patients, symptoms, and disease necessary to form strong knowledge representations. The nature of rare or atypical diseases makes them unreliable learning experiences for medical students and residents. Medical education programs must create the situations necessary for learning by taking advantage of various simulation-based learning techniques and applying the principles of deliberate practice early in medical training.[81]

REFERENCES

1. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving: Ten-year retrospective. *Evaluation & the Health Professions* 1990;13:5–36.
2. Neufeld VR, Norman GR, Feightner JW, Barrows HS. Clinical problem-solving by medical students: A cross-sectional and longitudinal analysis. *Medical Education* 1981;15:315–22.
3. Patel VL, Groen GJ. Knowledge based solution strategies in medical reasoning. *Cognitive Science* 1986;10:91–116.
4. Eva KW, Brooks LR, Norman GR. Forward reasoning as a hallmark of expertise in medicine: Logical, psychological, and phenomenological inconsistencies. *Advances in Psychological Research* 2002;8:41–69.
5. Bordage G, Lemieux M. Semantic structures and diagnostic thinking of experts and novices. *Academic Medicine* 1991;66:S70–2.
6. Barrows HS, Feltovich PJ. The clinical reasoning process. *Medical Education* 1987;21:86–91.
7. Charlin B, Boshuizen H, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Medical Education* 2007;41:1178–84.
8. Papa FJ, Shores JH, Meyer S. Effects of pattern matching, pattern discrimination, and experience in the development of diagnostic expertise. *Academic Medicine* 1990;65:S21–2.
9. Patel VL, Groen GJ, Scott HM. Biomedical knowledge in explanations of clinical problems by medical students. *Medical Education* 1988;22:398–406.
10. Croskerry P. Clinical cognition and diagnostic error: Applications of a dual process theory of reasoning. *Advances in Health Sciences Education: Theory & Practice* 2009;14:27–35.
11. Croskerry P. Cognitive forcing strategies in clinical decision making. *Annals of Emergency Medicine* 2003;41:110–20.
12. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine* 2003;78:775–80.
13. Norman GR, Brooks LR, Allen SW. Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1989;15:1166.
14. Brooks LR, Norman GR, Allen SW. Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General* 1991;120:278.
15. Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: The role of experience. *Medical Education* 2007;41:1140–5.
16. Galotti KM. Cognitive psychology: In *and out of the laboratory. First Canadian Edition*. Toronto: Nelson Education, 2010.
17. Kahneman D. *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux, 2011.
18. Stanovich KE, West RF. On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology* 2008;94:672.
19. Klein, JG. Five pitfalls in decisions about diagnosis and prescribing. *British Medical Journal* 2005;330:781.
20. Redelmeier DA. The cognitive psychology of missed diagnoses. *Annals of Internal Medicine* 2005;142:115–20.
21. Evans JSB. Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology* 2008;59:255–78.
22. Evans JSB, Elqayam S. Dual-processing explains base-rate neglect, but which dual-process theory and how? *Behavior and Brain Science* 2007;30:261–2.
23. De Neys W. Heuristic bias, conflict, and rationality in decision-making. In B Glatzeder (Ed.), *Towards a theory of thinking* (pp. 23–33). Berlin: Springer, 2010.
24. Wilson TD, Brekke N. Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin* 1994;116:117.
25. Stanovich KE, West RF. Individual differences in reasoning: Implicaitons for the reationality debate. *Behavior and Brain Research* 2000;23:645–726.
26. De Neys W. Dual processing in reasoning two systems but one reasoner. *Psychological Science* 2006;17:428–33.
27. Graber ML, Kissam S, Payne VL, Meyer AN, Sorensen A, Lenfestey N, et al. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Quality and Safety* 2012;21:535–57.
28. Sherbino J, Dore KL, Wood TJ, Young ME, Gaissmeier W, Krueger S, et al. The relationship between response time and diagnostic accuracy. *Academic Medicine* 2012;87:785–91.
29. Norman GR, Sherbino J, Dore K, Wood TJ, Young M, Gaissmeier W, et al. The etiology of diagnostic errors: A controlled trial of system 1 vs. system 2 reasoning. *Academic Medicine* in press.
30. Ilgen JS, Bowen JL, McIntyre LA, Banh KV, Barnes D., Coates WC, et al. The impact of instruction to use first impressions or directed search on candidate diagnostic performance and the utility of vignette-based assessment. *Academic Medicine* 2013;88:535–41.
31. Monteiro SM, Sherbino JD, Ilgen JS, Dore KL, Wood TJ, Young ME, et al. Disrupting diagnostic reasoning: The effect of interruptions on the diagnostic performance of residents and emergency physicians. *Manuscript under review.*
32. Sherbino J, Dore KL, Siu E, Norman GR. The effectiveness of cognitive forcing strategies to decrease diagnostic error: An exploratory study. *Teaching and Learning in Medicine* 2011;23:78–84.
33. Sherbino J, Kulasegaram K, Howey E, Norman G. Ineffectiveness of cognitive forcing strategies to reduce biases in diagnostic reasoning: a controlled trial. *Canadian Journal of Emergency Medicine* 2013;15:1–7.
34. LeBlanc VR, Norman GR, Brooks LR. Effect of a diagnostic suggestion on diagnostic accuracy and identification of clinical features. *Academic Medicine* 2001;76:S18–20.
35. LeBlanc VR, Brooks LR, Norman GR. Believing is seeing: The influence of a diagnostic hypothesis on the interpretation of clinical features. *Academic Medicine* 2002;77:S67–9.
36. Brooks LR, LeBlanc VR, Norman GR. On the difficulty of noticing obvious features in patient appearance. *Psychological Science* 2000;11:112–4.
37. Norman GR, Colle C, Brooks LR, Hatala R. A disadvantage for searching medical stimuli without a diagnostic hypothesis. *Cognition and Instruction* 2000;17:433–48.
38. Schon D. *The reflective practitioner: How professionals think in action*. New York: Basic Books, 1983.
39. Mamede S, Schmidt HG, Rikers RMJP, Penaforte JC, Coelho-Filho JM. Breaking down automaticity: Case ambiguity and the shift to reflective approaches in clinical reasoning. *Medical Education* 2007;41:1185–92.
40. Mamede S, Schmidt HG, Rikers RMJP, Penaforte JC, Coelho-Filho JM. Influence of perceived difficulty of cases on physicians' diagnostic reasoning. Influence of perceived difficulty of cases on physicians' diagnostic reasoning. *Academic Medicine* 2008;83:1210–6.
41. Mamede S, Schmidt H, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnosis. *Medical Education* 2008;42:468–75.
42. Mamede S, van Gog T, van den Berge K, Rikers RMJP, van Saase JICM, van Guldener C, et al. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *Journal of the American Medical Association* 2012;304:1198–203.
43. Mamede S, Schmidt HG. The structure of reflective practice in medicine. *Medical Education* 2004;38:1302–8.
44. Ericsson KA, Charness N. Expert performance: Its structure and acquisition. *American Psychologist* 1994;49:725–47.
45. Beilock SL, Carr TH, MacMahon C, Starkes JL. When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied* 2002;8:6–16.

46. Eva KW, Regehr G. Knowing when to look it up: A new conception of self-assessment ability. *Academic Medicine* 2007;82:S81–4.
47. Lieberman MD, Jarcho JM, Satpute AB. Evidence-based and intuition-based self-knowledge: An fMRI study. *Journal of Personality and Social Psychology* 2013;87:421–35.
48. Woolley A, Kostopoulou O. Clinical intuition in family medicine: More than first impressions. *Annals of Family Medicine* 2013;11: 60–6.
49. Friedman CPF, Gateti GG, Franz Tm, Murphy GC, Wolf FM, Heckerling PS, et al. Do doctors know when their dignoses are correct? Implications for decision support systems. *Journal of General Internal Medicine* 2005;20:334–39.
50. Kruglanski AW. Only one? The default interventionist perspective as a unimodel—Commentary on Evans Stanovich. *Perspectives on Psychological Science* 2013;8:242–7.
51. Thompson VA. Why it matters the implications of autonomous processes for dual process theories—Commentary on Evans Stanovich. *Perspectives on Psychological Science* 2013;8:253–6.
52. Stanovich KE. Why humans are (sometimes) less rational than other animals: Cognitive complexity and the axioms of rational choice. *Thinking & Reasoning* 2013;19:1–26.
53. Keren G. A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans Stanovich. *Perspectives on Psychological Science* 2013;8:257–62.
54. Osman M. A case study: Dual-process theories of higher cognition—Commentary on Evans Stanovich. *Perspectives on Psychological Science* 2013;8:248–52.
55. Osman M. An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin and Review* 2004;11:988–1010.
56. Smith JD, Minda JP. Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1998;24:1411.
57. Yonelinas AP. The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language* 2002;46:441–517.
58. Jacoby LL. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language* 1991;30:513–41.
59. Gobet F, Simon HA. Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology* 1996;31:1–40.
60. Loftus EF, Palmer JC. Reconstruction of automobile destruction. An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior* 1974;13:585–9.
61. Godden DR, Baddeley AD. Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology* 1975;66:325–31.
62. Smith SM, Vela E. Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin and Review* 2001;8:203–20.
63. Neath I. *Human memory: An introduction to research, data, and theory*. Pacific Grove, CA: Thomson Brooks/Cole, 1998.
64. Nosofsky RM, Zaki SR. Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science* 1998;9:247–55.
65. Wills AJ, Pothos EM. On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin* 2012;138:102.
66. Minda JP, Smith JD. Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2001;27:775.
67. Bordage G., Zacks R. The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes. *Medical Education* 1984;18:406–16.
68. Yonelinas AP, Jacoby LL. Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Psychology* 1994;48:516.
69. Hockley WE. The effects of environmental context on recognition memory and claims of remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2008;34:1412.
70. Aggleton JP, Brown MW. Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Science* 2008;10:455–63.
71. Larsen DP, Butler AC, Roediger HL III. Test-enhanced learning in medical education. *Medical Education* 2008;42:959–66.
72. Hatala RM, Brooks LR, Norman GR. Practice makes perfect: The critical role of mixed practice in the acquisition of ECG interpretation skills. *Advances in Health Sciences Education: Theory and Practice* 2003;8:17–26.
73. Chisholm CD, Weaver CS, Whenmouth L, Giles B. A task analysis of emergency physician activities in academic and community settings. *Annals of Emergency Medicine* 2011;58:117–22.
74. Graber ML, Gordon R, Franklin N. Reducing diagnostic errors in medicine: What's the goal? *Academic Medicine* 2002;77:981–92.
75. Eva KW, Norman GR. Heuristics and biases—A biased perspective on clinical reasoning. *Medical Education* 2005;39:870–2.
76. Croskerry P. The cognitive imperative: Thinking about how we think. *Academic Emergency Medicine* 2000;7:1223–31.
77. Zwaan L, Thijs A, Wagner C, van der Wal G, Timmermans DR. Relating faults in diagnostic reasoning with diagnostic errors and patient harm. *Academic Medicine* 2012;87:149–56.
78. Graber ML. Diagnostic error in internal medicine. *Archives of Internal Medicine* 2005;165:1493–9.
79. Eva KW. Diagnostic error in medical education: Where wrongs can make rights. *Advances in Health Sciences Education: Theory and Practice* 2009;14:71–81.
80. Ark T, Brooks LR, Eva K. Giving learners the best of both worlds: Do clinical teachers need to guard against teaching pattern recognition to novices? *Academic Medicine* 2006;81:405–9.
81. Norman GR. Medical education: Past, present, future. *Perspectives on Medical Education* 2012;1:6–14.