Canadian Medical Education Journal

Major Contribution/Research Article

Experienced physicians benefit from analyzing initial diagnostic hypotheses

Adam Bass,[1] Colin Geddes,[2] Bruce Wright,[1] Sylvain Coderre,[1] Remy Rikers,[3] Kevin McLaughlin[1]

[1]University of Calgary, Calgary, Alberta, Canada.
[2]Glasgow University, Glasgow, Scotland, UK.
[3]Erasmus University, Rotterdam, the Netherlands

Abstract

Background: Most incorrect diagnoses involve at least one cognitive error, of which premature closure is the most prevalent. While metacognitive strategies can mitigate premature closure in inexperienced learners, these are rarely studied in experienced physicians. Our objective here was to evaluate the effect of analytic information processing on diagnostic performance of nephrologists and nephrology residents.

Methods: We asked nine nephrologists and six nephrology residents at the University of Calgary and Glasgow University to diagnose ten nephrology cases. We provided presenting features along with contextual information, after which we asked for an initial diagnosis. We then primed participants to use either hypothetico-deductive reasoning or scheme-inductive reasoning to analyze the remaining case data and generate a final diagnosis.

Results: After analyzing initial hypotheses, both nephrologists and residents improved the accuracy of final diagnoses (31.1% vs. 65.6%, $p < 0.001$, and 40.0% vs. 70.0%, $p < 0.001$, respectively). We found a significant interaction between experience and analytic processing strategy ($p = 0.02$): nephrology residents had significantly increased odds of diagnostic success when using scheme-inductive reasoning (odds ratio [95% confidence interval] 5.69 [1.59, 20.33], $p = 0.07$), whereas the performance of experienced nephrologists did not differ between strategies (odds ratio 0.57 [0.23, 1.39], $p = 0.20$).

Discussion: Experienced nephrologists and nephrology residents can improve their performance by analyzing initial diagnostic hypotheses. The explanation of the interaction between experience and the effect of different reasoning strategies is unclear, but may relate to preferences in reasoning strategy, or the changes in knowledge structure with experience.

Correspondence: Dr. Kevin McLaughlin, Office of Undergraduate Medical Education, University of Calgary, Health Sciences Centre, 3330 Hospital Drive NW, Calgary, Alberta, Canada T2N 4N1; Tel: (403) 220 4252; E-mail: kmclaugh@ucalgary.ca

Introduction

Despite ongoing improvements in physician training and healthcare delivery, 10 – 15% of patients are still misdiagnosed.[1] While diagnostic errors often have multiple etiologies, faulty cognition is implicated in approximately three quarters of all errors.[2] So why are cognitive errors so frequent?

Diagnosing, and decision-making in general, involves two cognitive processes.[3-5] The first is automatic information processing, also referred to as intuition or pattern recognition. This requires minimal use of working memory, and involves rapid, subconscious processing of data that are largely contextual, to reach a single diagnosis or a short list of possible diagnoses. By contrast, analytic information processing involves conscious evaluation of case data and potential diagnoses by working memory until the best fitting diagnosis is selected. Diagnostic error may result from faults in either, or both, of these processes, but the most frequent explanation is "premature closure", where a final diagnosis is selected before sufficient information has been processed.[1]

If insufficient analytic processing causes diagnostic error then further analysis of diagnostic hypotheses might improve performance.[6] This metacognitive approach is supported by several studies demonstrating mitigation of premature closure and improved accuracy of final diagnoses.[7-11] But further analysis does not always improve decision making, and the psychology and medical education literature is replete with studies demonstrating that analysis can result in poorer diagnoses/decisions.[12-14] This equipoise suggests that the effect of analytic processing on diagnostic performance may be modified by other variables, such as the structure and complexity of the content area being studied, experience within this content area, and perhaps individual preferences of problem solving strategies.[5]

These potential effect modifiers are not independent. Complexity is inversely related to experience, and problem-solving preferences change with experience. Novice learners solve problems by first creating mental models from their underlying knowledge and perception of the problem, and then analyze each model until the best fit is selected.[15,16]

When studying problem-solving in medicine we typically refer to "models" as "diagnostic hypotheses", and replace "model-based reasoning" with "hypothetico-deductive reasoning". With experience, this form of reasoning becomes increasingly complex as improved knowledge allows for a larger number of more elaborate models to be generated, thus increasing cognitive load.[15,17] To counter this, more experienced learners abstract domain-specific rules and then form rule-based schemas to allow them to analyze problems without building models, or by building fewer models.[18-20] In theory, adopting "rule-based reasoning" (or "scheme-inductive reasoning") allows for more efficient diagnosing. Finally, with further experience the role of analytic processing is diminished and diagnoses are increasingly made by automatic processing alone.[21] Experienced physicians may be afforded the luxury of diagnosing without analysis due to improved perception and ability to generate accurate initial hypotheses, and/or because they have developed mental shortcuts, or "heuristics", whereby they omit or obscure the intermediary steps of analytic processing.[22,23]

Experienced physicians develop heuristics as these are more likely to help than hinder diagnostic reasoning.[24] Yet with increasing experience and reliance on automatic processing and heuristics, physicians are more susceptible to premature closure, are less willing to reconsider their diagnoses, and, consequently, become progressively error prone.[25-27] If analyzing an initial hypothesis improves performance of less experienced learners, this may also work for experienced physicians. But we should not assume that metacognitive strategies are universally beneficial: if we force experienced physicians to abandon their preferred reasoning strategies and process information like novices they may perform like novices![28]

In this study our objective was to evaluate the effect of analyzing an initial hypothesis on the diagnostic performance of practicing physicians. We predicted that if analysis is an effective metacognitive strategy then performance should improve with analysis. To test this hypothesis we manipulated the information processing strategies of two groups of physicians with differing degrees of experience – practicing nephrologists with ten or more years experience in

caring for patients with kidney diseases, and nephrology subspecialty residents who had completed their training in internal medicine and were receiving further training in nephrology. After providing our participants with a brief clinical presentation and contextual information, we asked them to commit to an initial diagnosis – knowing that this would encourage the use of heuristics and risk premature closure. We then asked them to reconsider their diagnoses in light of further information on the case, and evaluated the impact of further analysis on diagnostic performance. Cognizant of the different analytic processing strategies, we primed our participants to use either hypothetico-deductive reasoning or scheme-inductive reasoning, thus allowing us to compare the impact of each strategy on diagnostic performance.

Methods

Our study participants were nine practicing nephrologists with more than ten years of clinical experience in the care of nephrology patients, and six nephrology subspecialty residents at the University of Calgary or Glasgow University. The nephrology subspecialty residents had completed at least three years of training in internal medicine and were enrolled in nephrology subspecialty training programs of two year duration. The Conjoint Medical and Research Ethics Board at the University of Calgary granted ethical approval for the study and each participant provided informed consent prior to entry into the study.

We randomized our participants to receive one of two booklets containing ten identical nephrology cases in the same order. Each case began with primary data that included the presenting complaint – which could be a symptom, sign, or abnormal test result – along with the patient's age, gender, enabling conditions, and clinical setting. Based upon these data alone, we asked our participants to offer an initial diagnostic hypothesis. The booklets then differed in the reasoning instructions given for each case: if group 1 was given instructions to use hypothetico-deductive reasoning for case 1 then group 2 was instructed to use scheme-inductive reasoning (see descriptions of these instructions

below). Using computer-generated random numbers, we randomized the sequence of reasoning instructions, and each group answered five questions using each reasoning strategy. Cognizant that restricting the time available for information processing might increase the risk of premature closure, we did not limit the time available to complete problems. Figure 1 provides an overview of our study design.

Instructions for hypothetico-deductive reasoning

Instructions on hypothetico-deductive reasoning were similar to those used by Mamede et al.[10] After offering an initial diagnostic hypothesis participants were asked to list their differential diagnoses. Following the presentation of secondary data, they were asked to consider, for each diagnosis, features consistent with this diagnosis, features inconsistent with this diagnosis, and features that would have been expected if this were the correct diagnosis. We then asked them to rank diagnoses and give their final diagnosis.

Instructions for scheme-inductive reasoning

After offering an initial diagnostic hypothesis participants were asked to draw a diagnostic scheme that could be used to help diagnose the cause of this clinical presentation. We then asked them to use their diagnostic scheme to analyze secondary data and provide their final diagnosis.
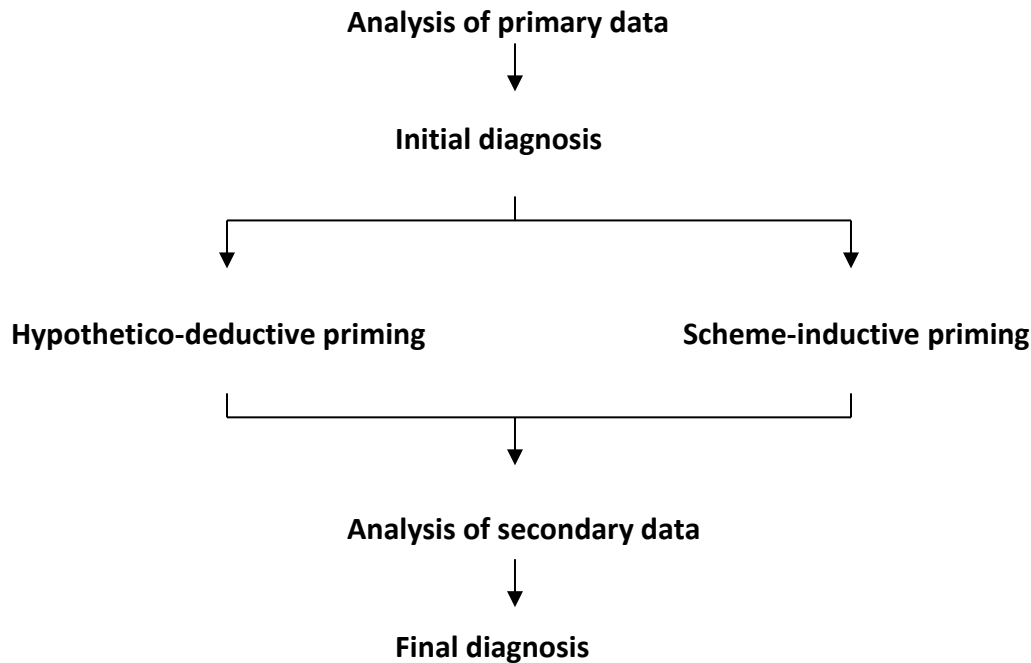
Evaluation of diagnostic performance

The cases used were based upon real patients. For each case the final diagnosis was either confirmed using the available gold standard, such as histology, or two experienced nephrologists (CG & KM) agreed upon a single best diagnosis. Participants' diagnoses were considered correct if they matched, or were synonymous with, the agreed-upon final diagnosis.

Statistical Analyses

We used Fisher's exact test to compare the proportion of correct diagnoses for different degrees of clinical experience (experienced nephrologist vs. nephrology resident), and reasoning strategies. We used McNemar's discordant pair analysis to compare the direction of change (incorrect to correct and vice versa) between initial and final diagnoses.

Figure 1. Overview of the study design.

Analysis of primary data

↓

Initial diagnosis

Hypothetico-deductive priming   **Scheme-inductive priming**

Analysis of secondary data

↓

Final diagnosis

We used multiple logistic regression to evaluate the effect of reasoning strategy and clinical experience on diagnostic performance. In our regression model we also considered two-variable interaction terms, and used backward elimination to remove non-significant variables from the model. We used STATA version 11.0 for our statistical analyses.

Results

The effect of analysis on diagnostic performance

The mean accuracy of initial diagnostic hypotheses was 34.7%, and this did not differ between experienced nephrologists and nephrology residents (p = 0.30). After analysis of the initial hypothesis, 35% of diagnoses were changed. Of these, the vast majority resulted in an incorrect initial hypothesis being changed to a correct final diagnosis rather than a correct initial hypothesis being changed to an incorrect final diagnosis (96% vs. 4%, p < 0.0001). The likelihood of changing an initial diagnostic hypothesis did not differ between nephrologists and residents (p = 0.70), or the information processing strategy used (p = 0.90).

After analyzing their initial diagnostic hypotheses, both nephrologists and residents improved the accuracy of their final diagnoses (31.1% vs. 65.6%, p <0.001, and 40.0% vs. 70.0%, p < 0.001, respectively). These data are shown in Figure 2.

Hypothetico-deductive reasoning vs. scheme-inductive reasoning

In our logistic regression model we found a significant interaction between clinical experience and the effect of different analytic processing strategies on diagnostic performance (p < 0.01). Stratifying our analysis by clinical experience, we found that nephrology residents had significantly increased odds of diagnostic success when using scheme-inductive reasoning as compared to hypothetico-deductive reasoning (odds ratio [95% confidence interval] 5.69 [1.59, 20.33], p < 0.01), whereas the performance of experienced nephrologists did not differ between strategies (odds ratio for scheme-inductive reasoning was 0.57 [0.23, 1.39], p = 0.20). Figure 3 shows the effect of the different analytic strategies on diagnostic performance.

Figure 2. The effect on diagnostic performance of analyzing an initial hypothesis
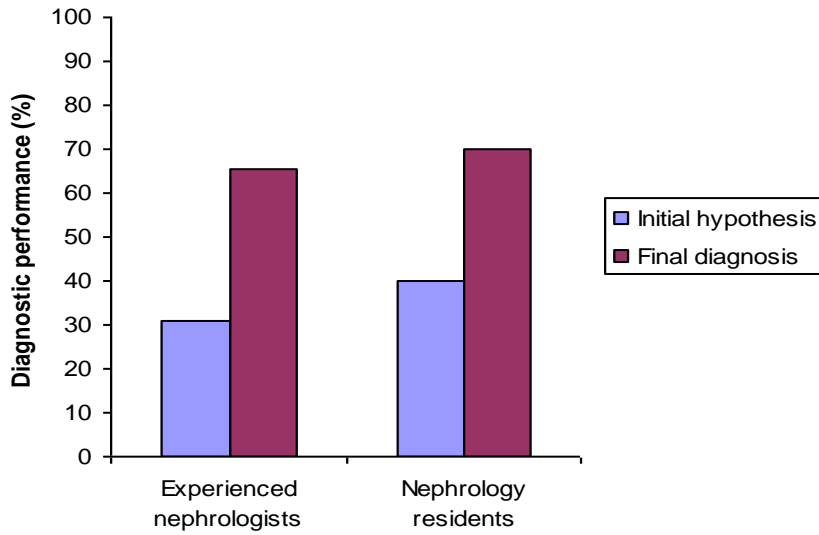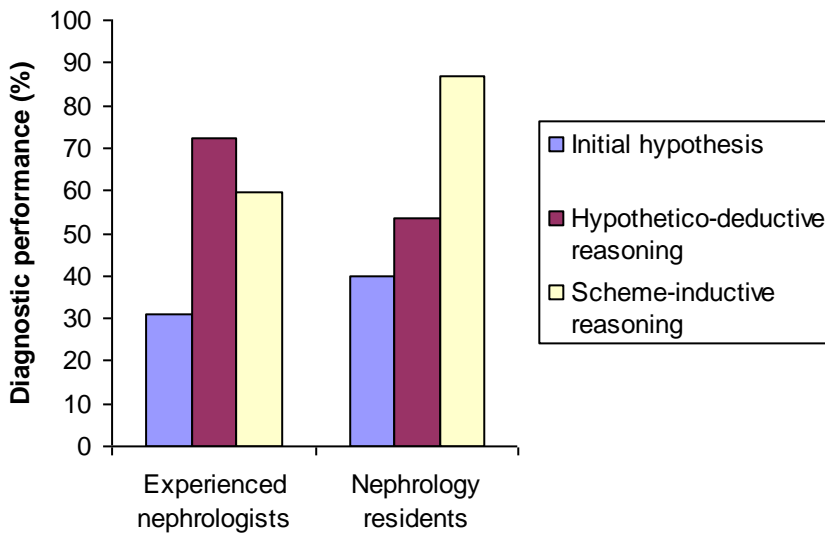

Figure 3. The effect on diagnostic performance of hypothetico-deductive reasoning vs. scheme-inductive reasoning

When we analyzed compliance with priming instructions we found that our participants always followed the priming instructions for hypothetico-deductive reasoning, but experienced nephrologists failed to generate a diagnostic scheme (or described domain-specific rules) for 70% of the cases where they were asked to do so. By comparison, the non-compliance rate with scheme-inductive reasoning priming conditions among nephrology residents was only 3%. We therefore repeated our analyses based upon actual, rather than intended, analytical processing strategy used. In this post-hoc analysis the effect of analytic reasoning strategy on diagnostic performance was unchanged: nephrology residents had an increased odds of diagnostic success using scheme-inductive reasoning (odds ratio 5.15 [1.44, 18.34], $p < 0.01$), while there was no difference in the performance of experienced nephrologists (odds ratio 0.65 [0.20, 2.09], $p = 0.50$).

Discussion

This study adds to the growing literature on the diagnostic benefits of analytic information processing. But rather than evaluating this approach in non-physicians or inexperienced trainees, our participants were physicians with considerable clinical experience – including those who are typically considered to be less flexible in their diagnostic reasoning strategies, and more prone to diagnostic error as a result of premature closure.[25-27] The major finding in our study was that both experienced nephrologists and residents receiving subspecialty training in nephrology improved their performance after analyzing their initial diagnostic hypotheses. Upon analysis, they rarely substituted correct hypotheses with wrong diagnoses, but frequently rejected incorrect hypotheses in favour of correct diagnoses. Overall, the diagnostic success rate approximately doubled after analyzing their initial hypotheses.

The benefits of analytic processing are unlikely to be explained by more information processing simply being better rather than less. There are too many examples of performance declining with over-analysis to accept this explanation.[12-14] In almost every field of study, experienced practitioners become progressively automated in their thoughts and actions (or "unconsciously competent"[29]), whereas analytic processing involves the use of working memory, where the logic of decisions can be examined, and their consequences anticipated. Studying experienced decision makers in many fields, Klein has concluded that experts – as distinct from simply experienced practitioners – use "recognition-primed decision making" where they consciously analyze the consequences of their initial decision, and then alter their decisions accordingly.[30] Concurring with this, Epstein used the term "mindful practice" to describe decision making by expert diagnosticians.[31] These descriptors imply that in the minds of experts cognitive processes complement each other rather than compete against each other, and that analysis is not better than intuition, or vice versa.[32] In our study we did not try to force our participants to abandon their heuristics in favour of analysis. Rather, we actively encouraged them to use these heuristics to generate an early hypothesis – after which we provided an opportunity to debias their reasoning process by analyzing their initial hypothesis in light of further data. Thus, we forced our participants to use recognition-primed decision making, seemingly to good effect.

So, if it pays to analyze diagnostic hypotheses, is there a better way of doing this? Based upon observational studies suggesting that analytic processing naturally evolves from hypothetico-deductive to scheme-inductive reasoning, we predicted that scheme-inductive reasoning would be a more effective strategy.[18-20] However, we found a significant interaction with experience: nephrology residents, but not experienced nephrologists, performed better when primed to use scheme-inductive reasoning.

The simple explanation for the interaction between experience and reasoning strategy is that residents prefer scheme-inductive reasoning, while experienced nephrologists prefer hypothetico-deductive reasoning. This is supported by the observation that most experienced nephrologists did not use scheme-inductive reasoning when primed to do so. Unfortunately our study was not designed to explain preferences in diagnostic reasoning, but we would speculate that each group prefers to do what they do on a daily basis. Nephrology residents, by

virtue of their mandatory supervision, are forced to use analytic processing for each case they see in order to justify their diagnosis to the attending physician. They have sufficient training in nephrology to learn the domain-specific rules, and, in addition to a limited number of illness scripts, their knowledge structure contains abridged causal networks, replete with biomedical knowledge, within which domain-specific rules are stored.[33,34] By comparison, experienced nephrologists are not forced to articulate the logic behind each diagnostic decision, and their knowledge structure is primarily composed of illness scripts, within which domain-specific rules have been encapsulated.[33-35] Failure to use domain-specific rules on a daily basis may make them less accessible to experienced nephrologists, such that may be easier for them to recall previously encountered diagnoses, rather than rules.

Our study has some limitations. The small number of participants limited our power to detect differences within and between groups related to the use of different analytical reasoning strategies which resulted in imprecision and wide confidence intervals in our results. We only studied one clinical domain, so we cannot extrapolate the benefits of analytic processing to other domains. Although complex, nephrology problems are usually highly structured, and the domain-specific rules are well established – which may exaggerate the diagnostic benefits of analytic processing.[5] We cannot assume that diagnosing in clinical domains where the rules are less clear, for example psychiatry, will also be enhanced by analytic processing. Although we observed a significant improvement in diagnostic performance after analytic processing, approximately one third of final diagnoses were incorrect. This was despite sufficient data being available on each case for the correct final diagnosis to be made. This suggests that other cognitive biases may have been operating – such as confirmation bias, where data supporting, rather than refuting, a diagnostic hypothesis, are over emphasized.[6] More than thirty cognitive biases have been identified, so no single cognitive strategy is a panacea.[6] There was no control group without a reasoning strategy and so

we cannot comment on how to participants would have performed without a reasoning strategy. This was deliberate in our design since without specific instructions, there would have been no way of accurately determining which reasoning strategy or combination of strategies participants would have employed. Finally, it is unclear why experienced nephrologists failed to use schemes in our study. While there are a number of possible explanations for this finding, it is impossible for us to accurately determine why this is the case. Further studies are needed to identify and mitigate other cognitive biases.

Implications for medical education

If the effect of a given information processing strategy on diagnostic performance may be modified by other variables, such as content area, experience, and individual preferences, making blanket recommendations on how to solve clinical problems is illogical. Recommendations on how to improve performance should target the cause(s) of poor performance, of which there are many. Our results suggest that for experienced physicians, as for novices, if the cognitive problem is premature closure, then analysis of initial hypotheses may be a cognitive solution. As to how these hypotheses should be analyzed, it appears that there is no consistently "better" analytic strategy, due to effect modification by other variables, such as clinical experience.

Conclusion

We found that the diagnostic performance of experienced nephrologists and nephrology residents improved when they combined heuristics with analysis of their initial diagnostic hypotheses. The effectiveness of the analytical strategy used, scheme-inductive vs. hypothetico-deductive reasoning, varied with experience – likely due to divergent preferences for reasoning strategies, which may be related to differences in underlying knowledge structure. Although encouraging, our results also highlight that there are multiple sources of cognitive error, so one cognitive solution will not solve all cognitive problems.

References

1. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med*. 2005;165:1493-1499.

2. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time. *JAMA*. 2003;289:2849-2856.

3. Sloman SA. The empirical case for two systems of reasoning. *Psychol Bull*. 1996;119:3-22.

4. Croskerry P. A universal model of diagnostic reasoning. *Acad Med*. 2009;84:1022-1028.

5. Pretz JE. Intuition versus analysis: strategy and experience in complex everyday problem solving. *Mem Cognit*. 2008;36:554-566.

6. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med.* 2003;78:775-780.

7. Ark TK, Brooks LR, Eva KW. Giving learners the best of both worlds: do clinical teachers need to guard against teaching pattern recognition to novices? *Acad Med*. 2006;81:405-409.

8. Ark TK, Brooks LR, Eva KW. The benefits of flexibility: the pedagogical value of instructions to adopt multifaceted diagnostic reasoning strategies. *Med Educ*. 2007;41:281-287.

9. McLaughlin K, Heemskerk L, Herman R, Ainslie M, Rikers RM, Schmidt HG. Initial diagnostic hypotheses bias analytic information processing in non-visual domains. *Med Educ*. 2008;42:496-502.

10. Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ*. 2008;42:468-475.

11. Coderre S, Wright B, McLaughlin K. To think is good: querying an initial hypothesis reduces diagnostic error in medical students. *Acad Med*. 2010;85:1125-1129.

12. Redelmeier DA, Shafir E. Medical decision making in situations that offer multiple alternatives. *JAMA*. 1995;273:302-305.

13. Kulatunga-Moruzi C, Brooks LR, Norman GR. Using comprehensive feature lists to bias medical diagnosis. *J Exp Psychol: Learn Mem Cog*. 2004;30:563-572.

14. Dijksterhuis A, Bos MW, Nordgren LF, van Baaren RB. On making the right choice: the deliberation-without-attention effect. *Science. 2006*;311:1005-1007.

15. Cherubini P, Mazzocco A. From models to rules: mechanization of reasoning as a way to cope with cognitive overloading in combinatorial problems. *Acta Psychol*. 2004;116:223-243.

16. Johnson-Laird PN. Deductive reasoning. *Ann Rev Psychol*. 1999;50:109-135.

17. Sweller J. Cognitive load during problem solving: effects on learning. *Cogn Sci*. 1988;12:257–85.

18. Cheng PW, Holyoak KJ. Pragmatic reasoning schemas. *Cogn Psychol*. 1985;17:391-416.

19. Cheng PW, Holyoak KJ. On the natural selection of reasoning theories. *Cognition*. 1989;33:285-313.

20. Chao SJ, Cheng PW. The emergence of inferential rules: the use of pragmatic reasoning schemas by preschoolers. *Cog Develop*. 2000;15:39-62.

21. Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: the role of experience. *Med Educ*. 2007;41:1140-1145.

22. Custers EJ, Boshuizen HP, Schmidt HG: The influence of medical expertise, case typicality, and illness script component on case processing and disease probability estimates. *Mem Cogn*. 1996;24:384-399.

23. Hill OW. Intuition: Inferential heuristic or epistemic mode? Imagination *Cog Person*. 1987–1988;7:137-154.

24. Eva KW, Norman GR. Heuristics and biases – a biased perspective on clinical reasoning. *Med Educ*. 2005;39:870-872.

25. Eva KW, Link CL, Lutfey KE, McKinlay JB. Swapping horses midstream: factors related to physicians' changing their minds about a diagnosis. *Acad Med*. 2010;85:1112-1117.

26. Eva KW, Cunnington JP. The difficulty with experience: does practice increase susceptibility to premature closure? *J Contin Educ Health Prof*. 2006;26:192-198.

27. Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. *Ann Inter Med*. 2006;142:260-273.

28. Ayers P, Chandler P, Sweller J. The expertise reversal effect. *Educ Psychol*. 2003;38:23-31.

29. http://en.wikipedia.org/wiki/Four_stages_of_competence

30. Klein G. *Sources of power: how people make decisions*. Cambridge, MA. The MIT Press, 1998.

31. Epstein RM. Mindful practice. *JAMA*. 1999;282:833-839.

32. Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ*. 2005;39:98-106.

33. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. *Acad Med*. 1990;65:611-621.

34. Boshuizen PA, Schmidt HG. On the role of biomedical knowledge in clinical reasoning by experts, intermediates, and novices. *Cogn Sci*. 1992;16:153-184.

35. de Bruin AB, Schmidt HG, Rikers RM. The role of basic science knowledge and clinical knowledge in diagnostic reasoning: a structural equation modeling approach. *Acad Med*. 2005;80:765-773.