# Context and Causal Structure Enhance Memory for Clinical Details

**Sarah L. Devantier (sdevanti@uwo.ca)**
Department of Psychology
The University of Western Ontario
London, ON N6A 5C2

**John Paul Minda (jpminda@uwo.ca)**
Department of Psychology
The University of Western Ontario
London, ON N6A 5C2

**Mark Goldszmidt (Mark.Goldszmidt@schulich.uwo.ca)**
Schulich School of Medicine & Dentistry,
The University of Western Ontario
London, ON N6A 5A5

**Aimee L. Skye (SkyeA2@macewan.ca)**
Department of Psychology
Grant McEwan University
Edmonton, AB T5J 4S2

**Nicole N. Woods (nikki.woods@utoronto.ca)**
Wilson Centre
University of Toronto
Toronto, Ontario M5G 2C4

## Abstract

Current research suggests a facilitatory role for basic biomedical knowledge in learning and retaining concepts related to medical diagnosis. But learning and performance may be influenced by other knowledge as well. Accordingly, we examined the effects of foundational knowledge beyond basic biomedical science on the learning and retention of medical information. Subjects were asked to study a handout detailing a percussive chest exam and several respiratory disorders. One group was presented with the information in a standard "textbook" format and the other group was presented with foundational knowledge about how sound travels though solids and liquids. The foundational knowledge group outperformed the control group in a memory task. We suggest that these subjects were able to create causal links between the information to be learned and the foundational knowledge which made the critical information more memorable.

**Keywords**

Medical Reasoning, Memory, Foundational Knowledge, Causal Knowledge

Learning to perform a clinical exam or to make a diagnosis takes many years of training and practical experience. As students progress through medical school, they acquire skills first through formal classroom experience and later via direct clinical experience. In either case, acquiring new information can benefit from a rich knowledge base, especially to the extent that the knowledge base allows the learner to create causal links among the various features of new information (Murphy & Medin, 1985).

Researchers in cognition have shown that a sensitivity to causal information is beneficial when learning new information, and this seems to be true across many domains. For example, researchers have noted that clinical psychologists are likely to weight and subsequently recall perceived causal symptoms more heavily during a diagnosis (Kim & Ahn, 2001, 2002) or in a categorization task (Ahn & Luhmann, 2005) than features that are not perceived as causal. Similarly, causal features are considered to be more central and more important than non-causal features when categorizing novel stimuli or making inductions about a category (Ahn, 1998; Murphy & Medin, 1985; Rehder & Hastie, 2001; Sloman, Love, & Ahn, 1998); this is particularly true for the initial feature in a causal chain (Rehder & Kim, 2006).

The role of background and causal knowledge in medical reasoning is also well established. For example, Patel and collogues (Patel & Groen, 1986) suggest that physicians who arrive a correct diagnosis are more likely to have done so via causal chains and forward reasoning. Other work has examined the relative differences between medical education programs that either emphasized foundational learning or problem-based learning. Patel and collogues found support for better performance by students in what they termed "Conventional Curriculum" programs that emphasize the acquisition of biomedical knowledge (Patel, Groen, & Norman, 1991) than those in a "Problem-based Curriculum," though all students studied did seem to rely on foundational, biomedical knowledge.

While Patel's work is informative, it is largely descriptive; reachers have recently turned to more tightly-controlled psychological investigations. For example, Woods, Brooks, and Norman (2005) compared diagnostic performance when subjects (psychology undergraduate students) learned the probabilities of features given a disorder presented in tabular format with performance when subjects learned the same information in the context of biomedical causal knowledge. Subjects were asked to learn a set of diagnostic materials and were tested twice. On immediate testing, subjects in both the probability and biomedical knowledge conditions learned the material equally well, however when tested again one week later, the probability learners' scores had significantly decreased while causal knowledge learners' scores remained the same (Woods et al., 2005). The authors concluded that while probability information helped participants learn information in the short-term, it was not as successful a strategy for memory retention as causal biomedical knowledge. However, other research (Norman, Eva, Brooks, & Hamstra, 2006) suggests that during diagnosis novices rely mainly on causal knowledge (the basic mechanisms of disease) while experts rely more on inferences drawn from causal knowledge. This prompts the question: Is biomedical knowledge only useful to novices?

Further research examined whether experts use biomedical knowledge at all, and if so, precisely when they do so. One suggestion is that doctors move through three types of mental representations as they progress from novices to experts: from (1) biomedical information alone; to (2) illness scripts; and finally to (3) exemplars from their own experience (Norman, 2005; Schmidt, Norman, & Boshuizen, 1990; Schmidt & Boshuizen, 1993). These studies do indicate, however, that while biomedical information becomes encapsulated into higher-level causal models, all layers of information are always available, and biomedical knowledge can be accessed when higher-level concepts cannot solve a given problem (Schmidt & Boshuizen, 1993; Verkoeijen, Rikers, Schmidt, van de Wiel, & Kooman, 2004). Expert clinicians may have superior knowledge of basic science concepts, but only use it explicitly in difficult or complex cases (Norman, 2005). When asked to diagnose or describe patients, experts tend to provide course, diagnosis, frequency and treatment information, whereas novices tend to mostly provide biomedical information (Custers, Boshuizen, & Schmidt, 1998). When reaching a correct diagnosis in a complex case or in an area they are not proficient in, however, experts are more likely to explicitly use knowledge of disease mechanisms, or biomedical knowledge, to reason their way through the symptoms. In other words, when experience is insufficient, experts will use biomedical knowledge as a fall back (Regehr, Cline, Norman, & Brooks, 1994).

With this in mind, Woods and colleagues (2006) examined the relationship between time constraints and reliance on biomedical knowledge. They found that subjects who learned via causal explanations performed better than a control group when time was limited and were better able to recognize encapsulated and novel consistent words. More recently, Woods, Brooks and Norman (2007) tested the retention of diagnostic information and the amount of reliance on causal information. Subjects learned artificial diseases, with or without a causal explanation for each symptom. Subjects who received causal biomedical knowledge performed better on a delayed test than subjects who studied only a list of features for each disease. In a second experiment, subjects again learned four artificial diseases, each with three non-causal features and one causal chain. When tested immediately, subjects assigned equal weighting to causal and non-causal alternatives, but began to consider the causal alternative as more likely on the delayed test. In other words, with a delay, subjects relied on the causal alternative as the most probable, because its causality made it the most memorable. In terms of the encapsulation theory, symptoms linked by causal explanations were encapsulated while non-causal symptoms were not.

Rikers and colleagues (Rikers, Loyens, Winkel, Schmidt, & Sins, 2005; Rikers, Schmidt, & Moulaert, 2005) agree that biomedical knowledge becomes encapsulated within clinical knowledge through applying such knowledge during patient interactions. They suggest that while expert doctors may not explicitly state biomedical knowledge when diagnosing a patient, such knowledge supports diagnostic decisions. Rikers et al. demonstrated this using a priming paradigm. Subjects first read a short description of a hypothetical patient, followed by a mask and a target word. In one experiment (Rikers, Loyens, et al., 2005) the word was either a diagnostic or biomedical term, or a non-word, and subjects simply had to determine if the target was a word or non-word as quickly as possible. In a second experiment (Rikers, Schmidt, & Moulaert, 2005), subjects were shown a diagnostic or biomedical term and they had to indicate whether the item was related or unrelated to the previously read case. The results of both experiments indicated (through faster reaction times) that biomedical as well as diagnostic information was active for experienced doctors when reading a patient profile.

The results of all of these studies suggest a prominent role for biomedical knowledge as a facilitator in creating causal links between features and concepts. However, experienced physicians (and successful students) likely rely on foundational knowledge outside of the purely biomedical field. For example, when learning about a percussive chest exam, students are likely to be taught about the technique for doing the exams, the variety of outcomes from the exam, and the associated diseases and conditions that underlie these different sounds. But another way to think about the exam is in the context of the acoustical properties of sound traveling through solids, liquids or gas. In some cases, the sound can be muted by the presence of liquid in the lung space, just as sound might be muted when travelling through any liquid medium. Students with foundational knowledge in basic acoustics and physics might learn the percussive chest exam more readily than students whose foundational knowledge is purely biomedical in nature.

## Experiment 1A

We designed an experiment to test this prediction. We examined the ability of two groups of subjects to learn material related the percussive chest exam and several related diseases of the lung. One group, the Foundational Knowledge (FK) group, learned the material in the context of acoustical knowledge. The other group, the No Foundational Knowledge (NFK) group, served as the control group and was given the information without the acoustic information as they would typically learn it in a class in medical school. Both groups then completed a set of test items (immediately and after a one week delay). We predicted that the FK group would perform better than the NFK Group on test items that were directly related to the chest exam.

### Method

**Subjects** Subjects were recruited from the Undergraduate Psychology Research Pool at the University of Western Ontario and awarded course credit for their participation. Additional subjects were recruited from the undergraduate population at Grant McEwan College. A total of 88 subjects, 63 females and 25 males, aged 18-40 (M = 20.4, SD = 5.55) were tested. Subjects were randomly assigned into the one of the two learning groups. All of our subjects were collected from a first year psychology class and were not expected to have any previous medical knowledge that would help them on any of the tests. Twenty-two subjects were removed from the final analysis: Sixteen subjects (5 from the NFK group and 11 from the FK group) were removed from the final analysis for failing to return for the second test, and 5 subjects (4 from the NFK group and 1 from the FK group) were removed because they had an average score of less than .40 on the immediate test. One subject was removed from the NFK group because he had a background in university-level physics.

**Materials** Subjects in the No Foundational Knowledge group were given an information booklet titled “An Introduction to Diseases Involving the Lung.” The information in the booklet was designed by one of the authors (Goldszmidt) and involved the explanations for the causes and symptoms of three lung diseases (chronic obstructive pulmonary disease, pneumonia and pleural effusion) as well as a description on how to do a basic physical exam of the lungs by listening to breath sounds, percussing the lung and feeling for tactile fremitus. The Foundational Knowledge group had an extra section in their booklet titled “The Basic Physics of Lung Sounds” that was not included in the control group booklet. This extra section provided the foundational knowledge that was expected to make the earlier information presented more coherent and potentially more memorable. The information included physics rules and explanations for the results of each of the three parts to the physical exam of the lung. For example, all subjects read about the exam and how it would sound in a patients with chronic obstructive pulmonary disease (COPD). Subjects in both groups were given information about the disease, and the symptoms, and were told what examination outcomes would be indicative of COPD. However, subjects in the FK group were also told to *“imagine tapping in the same way on a drum. If it was a good drum, you would get a nice hollow or hyperresonant sound. This is like what happens in COPD with emphysema; a lot of the normal lung tissues are destroyed and so the sound is not dampened much at all.”*

Subjects’ retention was measured via two multiple choice quizzes that were administered immediately after study and one week later. We included fifteen critical test items and ten multiple choice control items. Test items were questions that required the participant to make a diagnosis based on exam results, or to predict a symptom given a diagnosis. As such, we expected that performance on these items would be aided by the foundational knowledge presented in the last section of the treatment group’s information booklet. The ten control items tested knowledge learned from the information booklet that was not aided by the presence of foundational knowledge. For instance, control items include *“Which is NOT a usual factor that places a person at higher risk for pneumonia?”* and *“The separation between the two outside layers of the lung is referred to as:”* where answering correctly to either question is not helped by knowing the physics information presented to the FK group. The only differences between the immediate test and the delayed test were question order, answer order, and patient names in the case studies.

**Procedure** Subjects were tested in small groups of up to six individuals in two sessions each in a quiet laboratory setting. During the first session, informed consent was obtained from each subject, and information on subjects’ English proficiency was also collected to control for possible differences in language ability and understanding. Subjects were then asked to read and study the paper information booklet. Specifically, they were instructed to study the booklet to the best of their ability because the information would not be presented again before any of the following tests and they were told to give back the booklet to the experimenter only after they felt confident they knew the information. There was no time limit on how long subjects could study the booklet or on either test; however, most took between 20 to 30 minutes to study the material and 7 to 15 minutes to do each test. The time it took each subject to get through each session was recorded to look for outliers or individuals who were having a harder time with the information; however, no outliers were observed. After a subject returned the information booklet to the experimenter, the immediate test was administered. The items were presented via an online survey application created in one of the author’s lab (Minda). After completing the test, subject booked their second session for exactly one week later. During the second session, subjects were given only the computerized test without being presented with the information booklet. After completing the experiment, subjects were de-

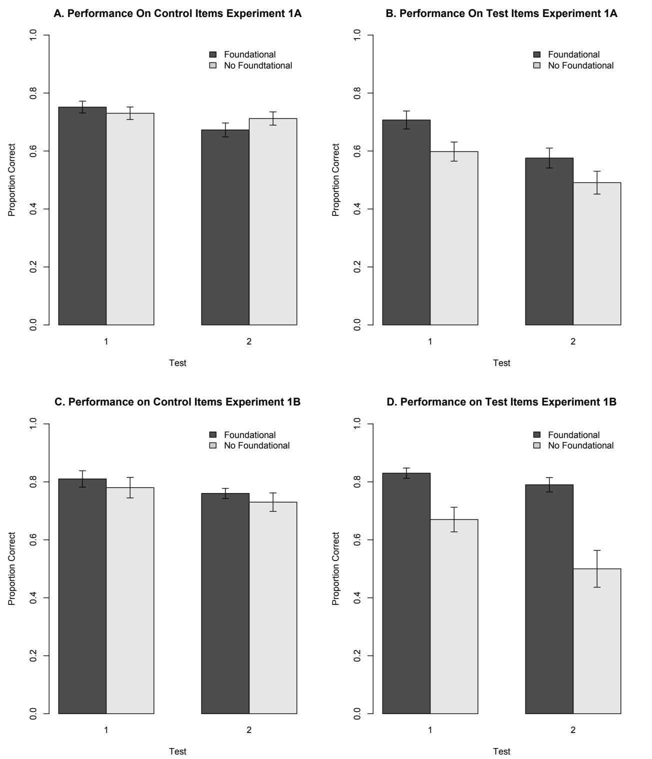


Figure 1: Panel A shows performance on control items when tested immediately and 1-week-later in experiment 1A. No significant difference was observed. Panel B shows the performance on test items: as expected, the FK group performed significantly better at Time 1, and showed a similar trend at Time 2. Panels C and D show the same information for experiment 1B. Again, there was no significant difference on control items, while the FK group performed better than the NFK group at both Time 1 and Time 2. *Note:* error bars denote *SEM*.

briefed. The average running time for completing the first session was 42 minutes and the average running time for completing the second session was 10 minutes.

## Results

**Learning Analysis** Each item was scored, and we calculated a proportion correct for each subject overall, as well as for control items and test items separately. Prior to conducting the main analysis, we conducted an item analysis. For each item, we calculated the correlation between subjects' responses to that item and subjects' overall performance. Our average correlation between item and performance was .34 (a range from 0.02 to .54), indicating that most of our individual items tracked overall competence in a reliable way.

We then averaged across subjects to obtain the average proportion correct for subjects in the Foundational Knowledge (FK) and the No Foundational Knowledge (NFK) groups on Test 1 and Test 2. The resulting averages are displayed in Figure 1.

As can been seen in the top panel of Figure 1, there was no almost difference between the two groups on the control items. We entered the scores into a mixed factorial ANOVA with Condition (FK/NFK) as a between subjects factor and Test (Test 1 / Test 2) as a within subjects factor. We failed to find any significant differences between the conditions $F(1,126) = 0.1062, p = .75$, indicating that the control items were not affected by the addition of foundational knowledge and that both groups learned the information equally. As well, memory of the control items did not decrease with time and there was no significant difference between the immediate and delayed control item test scores, $F(1,126) = 0.8484, p = .36$.

In accordance with our hypothesis, panel B in Figure 1 shows an advantage for the FK group over the NFK group on the test items. We entered the scores into a mixed factorial ANOVA with Condition (FK/NFK) as a between subjects factor and Test (Test 1 / Test 2) as a within subjects factor. We found a significant main effect of condition (performance was better for the Foundational Knowledge group) on the test items, $F(1,126) = 8.2056, p = .005$. There was also an effect of Test (performance on Test 1 was better than on Test 2) $F(1,126) = 4.3883, p = .038$; however, there was no interaction $F(1,126) = 0.1256, p = .72$.

Planned comparison t-tests (using a one-tailed two-sample t-test) on Test items indicated a significant difference at Time 1 ($t(64) = 2.4045, p = .0096$), and a trend towards a difference at Time 2 ($t(64) = 1.6246, p = .055$). In both cases, as expected, the FK group outperformed the NFK group.

# Experiment 1B

Experiment 1B was a direct replication of 1A. While considering the results of the first study there was a suggestion that the FK group received information that was more imagable (using the metaphor of a drum to describe the lungs), and therefore the FK group could be utilizing dual coding while the NFK could use only verbal coding (see, for example, Clark & Paivio, 1991). This prompted us to examine our testing materials for differences, and a single reference to the lungs as a "drum" was removed from the NFK information package.

## Method

**Subjects** Seventeen subjects (8 females) aged 18-21 ($M = 20.5, SD = 1.68$) were recruited from the University of Western Ontario. One subject from the NFK group was removed because they had an average score of less than .40 on the immediate test.

**Materials and Procedure** The same test items and testing procedures were used, except that all data were collected in paper and pencil format, rather than online, and as mentioned above, a single reference to the lungs as a drum was removed from the NFK information package.

## Results

As can be seen in Figure 1, Panel C, both groups had good performance on the control items and there were no differences in performance by either group $F(1,26) = 2.592, p = .120$, indicating that the control items were not affected by the addition of causal knowledge. As well, memory of the control items did not decrease with time and there was no significant difference between the immediate and delayed control item test scores, $F(1,26) = 0.0006, p = .980$. With regards to the test items (Figure 1, Panel D), a significant main effect of condition (performance was better for the Foundational Knowledge group) on the test items was found, $F(1,26) = 32.193, p < .0001$. There was no main effect of Test, $F(1,126)1.433, p = .242$, and no interaction was found $F(1,126) = 2.756, p = .109$. Planned comparison t-tests on Test items indicated a significantly superior performance of the FK group at both Time 1 ($t(14)3.498, p = .006$), and Time 2 ($t(14) = 4.341, p = .002$).

## Discussion

The results of this study suggest that our subjects were able to utilize foundational knowledge when learning new information. The primary evidence for this was the superior performance by the FK group on the critical test items. We suggest that the foundational physics knowledge (the tutorial on acoustics) allowed for the creation of a causal chain to guide the new knowledge acquisition. For example, students in both the FK and NFK groups know that a normal lung sounds hollow (when tapped from the back), and dull if the patient has pneumonia. However, students in the FK condition can create the causal chain: A drum filled with water sounds dull; when a person has pneumonia their lungs are filled with fluid; therefore, when a person has pneumonia, percussing their lungs must result in a dull noise. In other words, they now know *why* the exam produced the sound that it does. And this additional, causal knowledge resulted in better performance on items that were related to the exam itself. This additional knowledge did not relate (causally or otherwise) to the control items and so there was no advantage by the FK group.

While our results were generally in line with our hypothesis, we also expected the FK advantage to persist, and possibly increase over time as has been shown in other research (Woods et al., 2005, 2007). The FK advantage seemed to persist in Study 1A, though the *p* value was not quite significant at the .05 level at the second test. However, the FK advantage was very strong in both the first and second tests in Study 1B, and we did not find any evidence of an interaction between condition and time, suggesting that our results were robust to the one week delay. Study 1B thus replicated our main finding and clarified Study 1As marginal effect of knowledge on the second test.

This study has two main implications. Firstly, if replicated in a larger sample with first year medical students, it would suggest that the standard physical exam textbooks should include this type of information for all students. Secondly, similar to prior work, it would suggest that one of the criteria for choosing to integrate a basic science topic into the curriculum should relate to the extent to which the topic supports the development in students of causal explanations for clinical findings. This is particularly salient with regards to this study because, unlike the prior studies, this is the first of its kind to consider physics (acoustics) as a basic science relevant to the teaching of medicine. Moreover, supporting causal knowledge does not appear to have been a criterion considered in either the Carnegie report on the future of medical education nor in a recent study related to the role of the basic sciences in medical education (Finnerty et al., 2010; Irby, Cooke, & O'Brien, 2010). In the latter study (Finnerty et al., 2010), the value and role of the sciences was felt to relate to having a fundamental knowledge of the human body and for developing effective thinking skills but not for developing what Murphy and Medin (1985) would have called conceptual coherence (i.e. causal knowledge).

The only difference methodologically between experiments 1A and 1B was a reference to the lungs as a drum. While we believe the major contribution of this study is the usefulness of foundational knowledge outside of biomedical information when teaching medical students, this does raise the possibility that imagery or dual coding did play a role in how memorable the information was. Further studies will be required to fully expolore this idea.

In any case, these data do support the general claim that foundational knowledge (even outside of basic biomedical science) can and does help with the learning and retention of diagnostic information.

## Acknowledgments

This research was supported by a Research Developments Initiative Grant from the Social Science and Humanities Research Council of Canada to the second author.

Correspondence should be addressed to John Paul Minda at the Department of Psychology, The University of Western Ontario, London, ON N6A 5C2, CANADA, by email at `jpminda@uwo.ca`, or at `http://www.mindalab.com/`.

## References

Ahn, W.-K. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition*, *69*, 135-178.

Ahn, W.-K., & Luhmann, C. C. (2005). Demystifying theory-based categorization. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time. Carnegie Mellon symposia on cognition* (p. 277-300). Mahwah, NJ: Lawrence Erlbaum Associates.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149-210.

Custers, E. J. F. M., Boshuizen, H. P. A., & Schmidt, K. G. (1998). The role of illness scripts in the development of medical diagnostic expertse: Results from an interview study. *Cognition and Instruction*, *16*(4), 367-398.

Finnerty, E. P., Chauvin, S., Bonaminio, G., Andrews, M., Carroll, R. G., & Pangaro, L. N. (2010). Flexner revisted: The role and value of the basic sciences in medical education. *Academic Medicine*, *85*(2), 349-355.

Irby, D. M., Cooke, M., & O'Brien, B. C. (2010). Calls for reform of medical education by the Carnegie Foundation for the advancement of teaching: 1910 and 2010. *Academic Medicine*, *85*(2), 220-227.

Kim, N. S., & Ahn, W.-K. (2001). Theory-based reasoning in clinical psychologists. In *Proceedings of the twenty-third annual meeting of the cognitive science society.* Edinburgh, Scotland: Lawrence Erlbaum Associates.

Kim, N. S., & Ahn, W.-K. (2002). The influence of naive causal theories on lay concepts of mental illness. *American Journal of Psychology*, *115*, 33-65.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.

Norman, G. R. (2005). Research in clinical reasoning: Past history and current trends. *Medical Education*, *39*(4), 418-427.

Norman, G. R., Eva, K. W., Brooks, L. R., & Hamstra, S. (2006). Expertise in medicine and surgery. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (p. 339-353). Cambridge (U.K.): Cambridge University Press.

Patel, V. L., & Groen, G. J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91-116.

Patel, V. L., Groen, G. J., & Norman, G. R. (1991). Effects of conventional and problem-based medical curricula on problem solving. *Academic Medicine*, *66*, 380-389.

Regehr, G., Cline, J., Norman, G. R., & Brooks, L. R. (1994). Effect of processing strategy on diagnostic skiill in dermatology. *Academic Medicine*, *69*(10), S34-S36.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *130*(3), 323-360.

Rehder, B., & Kim, S. (2006). How causal knowledge affects classification: A generative theory of categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32*(4), 659-683.

Rikers, R. M. J. P., Loyens, S., Winkel, W., Schmidt, H. G., & Sins, P. H. M. (2005). The role of biomedical knowledge in clinical reasoning: A lexical decision study. *Academic Medicine*, *80*(10), 945-949.

Rikers, R. M. J. P., Schmidt, H. G., & Moulaert, V. (2005). Biomedical knowledge: Encapsulated or two worlds apart? *Applied Cognitive Psychology*, *19*, 223-231.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*(3), 205-221.

Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. A. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine*, *65*(10), 611-621.

Sloman, S. A., Love, B. C., & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*(2), 189-228.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., Schmidt, H. G., van de Wiel, M. W. J., & Kooman, J. P. (2004). Case representation by medical experts, intermediates and novices for laboratory data presented with or without a clinical context. *Medical Education*, *38*, 617-627.

Woods, N. N., Brooks, L. R., & Norman, G. R. (2005). The value of basic science in clinical diagnosis: Creating coherence among signs and symptoms. *Medical Education*, *40*, 973-979.

Woods, N. N., Brooks, L. R., & Norman, G. R. (2007). The role of biomedical knowledge in diagnosis of difficult clinical cases. *Advances in Health Sciences Education*, *12*(4), 417-426.

Woods, N. N., Howey, E. H. A., Brooks, L. R., & Norman, G. R. (2006). Speed kills? Speed, accuracy, encapsulations and causal understanding. *Medical Education*, *40*, 107-112.