

## Expanding the basic science debate: the role of physics knowledge in interpreting clinical findings

Mark Goldszmidt · John Paul Minda · Sarah L. Devantier ·  
Aimee L. Skye · Nicole N. Woods

Received: 10 June 2011 / Accepted: 29 September 2011 / Published online: 15 October 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** Current research suggests a role for biomedical knowledge in learning and retaining concepts related to medical diagnosis. However, learning may be influenced by other, non-biomedical knowledge. We explored this idea using an experimental design and examined the effects of causal knowledge on the learning, retention, and interpretation of medical information. Participants studied a handout about several respiratory disorders and how to interpret respiratory exam findings. The control group received the information in standard “textbook” format and the experimental group was presented with the same information as well as a causal explanation about how sound travels through lungs in both the normal and disease states. Comprehension and memory of the information was evaluated with a multiple-choice exam. Several questions that were not related to the causal knowledge served as control items. Questions related to the interpretation of physical exam findings served as the critical test items. The experimental group outperformed the control group on the critical test items, and our study shows that a causal explanation can improve a student’s memory for interpreting clinical details. We suggest an expansion of which basic sciences are considered fundamental to medical education.

**Keywords** Causal knowledge · Learning · Retention · Basic science

---

MG and JPM are both lead authors on this paper. Order was determined alphabetically.

---

M. Goldszmidt (✉)

Schulich School of Medicine and Dentistry, The University of Western Ontario,  
Room 115, Health Sciences Addition, London, ON N6A 5C1, Canada  
e-mail: Mark.Goldszmidt@schulich.uwo.ca

J. P. Minda (✉) · S. L. Devantier

Department of Psychology, The University of Western Ontario, London, ON N6A 5C2, Canada  
e-mail: jpminda@uwo.ca  
URL: <http://www.mindalab.com>

A. L. Skye

Department of Psychology, Grant McEwan University, Edmonton, AB, Canada

N. N. Woods

Wilson Centre, University of Toronto, Toronto, ON, Canada

## Introduction

One of the challenges facing medical schools is that of determining curricular content [The Association of Faculties of Medicine of Canada (AFMC 2009)]. Curriculum developers and medical educators are constantly being pushed to maximize relevance and eliminate redundancy, while at the same time finding room in the curriculum to add content to help students achieve the many new and emerging competencies (AFMC 2009). This has caused many to question the relevance and role of the basic sciences in the undergraduate curriculum (Finnerty et al. 2010; Irby et al. 2010). And while there appears to be consensus on the need to include basic science teaching in the curriculum, less clear are the answers to the questions of which aspects of which sciences to include and when, where and how to teach them (Finnerty et al. 2010; Irby et al. 2010; Woods et al. 2007). According to the recent Carnegie foundation report, one method that needs further exploration is the integration of basic and clinical teaching (Irby et al. 2010). The purpose of the present study was to explore a method of achieving such integration using physics principles and instruction on the interpretation of findings from a respiratory physical exam. In order to understand the basis of this study, it is first necessary to examine some of the research on causal knowledge.

Research in cognition has shown that causal information—information that provides an explanation for how things work or how they are related—can be beneficial for retaining and applying newly learned information (Murphy and Medin 1985). This has since been shown to be true across many domains, including medicine (Keil 2006; Rehder 2010).

Several studies have shown the importance of causal information for the retention and application of newly learned information in novices. For example, Woods, Brooks, and Norman (Woods et al. 2007) tested the retention of diagnostic information and the amount of reliance on causal information. Participants who learned a causal biomedical explanation for diagnostic features performed better on a delayed test than participants who studied only a list of features for each disease. In a related experiment, when taught both causal and non-causal features of diseases, participants' appeared to rely more heavily on causal features in a delayed diagnostic test. Similar findings of delayed enhanced retention have also been found when participants were tested on diagnostic performance after learning biomedical causal knowledge or the statistical probabilities of features (Woods et al. 2005). Knowledge of causal explanations has also been shown to enhance performance of feature list learning when time constraints were applied (Woods et al. 2006). Moreover, it should be noted that in these studies, as in many others, the important factor in determining the value of biomedical knowledge seemed to relate to the clarity, plausibility, and stability of the causal explanations, not depth of understanding.

With regards to expert use of causal knowledge, research by Patel and colleagues suggests that physicians who arrive at a correct diagnosis are more likely to have done so via causal chains and forward reasoning (Patel and Groen 1986). Other research 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 (Norman et al. 2006). Further research has suggested that doctors move through three types of mental representations as they progress from novices to experts: from biomedical information alone; to illness scripts; and finally to exemplars from their own experience (Norman 2005; Schmidt and Boshuizen 1993; Schmidt et al. 1990). These studies do indicate, however, that while biomedical information becomes encapsulated into higher-level causal models, all layers of information are always available (Schmidt and Boshuizen 1993; Verhoeijen et al. 2004). One study has demonstrated that experts, when confronted

with very difficult cases, are more likely than intermediates (first and second year residents) to invoke basic science causal mechanisms. (Norman et al. 1994).

While results from these studies suggest a prominent role for the basic sciences as a facilitator in creating causal links between clinical features and concepts, what they do not answer is which sciences to include and when and how to teach them in an integrated fashion (Woods et al. 2007). For example, one of the basic-science topics that is not often taught is the physics of acoustics, despite its considerable relevance to learning to interpret the findings from numerous physical exams, particularly the respiratory exam. During this exam, using a combination of percussion of the thorax, feeling for fremitus (vibrations made by a patient's voice and transformed by the underlying lung tissues) and listening to breath sounds, a physician tries to determine the status of the underlying lung. While standard textbooks on the physical exam, such as Bates Guide to Physical Exam (Bickley et al. 2007) do provide students with tables listing what happens to the various exam findings in different disease states, we have not found one that provides a causal explanation for these findings. Causal explanations have occasionally been described in the literature (Weslby and Earis 2001).

The purpose of our study was to test the hypothesis that teaching students, in an integrated fashion, the basics of acoustics as it relates to the respiratory exam will enhance their ability to interpret the findings from a physical exam in various medical conditions. Based on the prior research on causal knowledge, we also hypothesized that retention of knowledge over time would be superior for students taught in this way.

## Study

### Method

Human research ethics approval was obtained from the University of Western Ontario and from Grant McEwan University.

### *Participants*

Participants were recruited from the Undergraduate Psychology Research Pool at the University of Western Ontario (UWO) and Grant MacEwan University (GMU). This population was chosen to ensure that they would not have had prior exposure to the respiratory physical exam. Participants were also excluded if they had a background in physics. A total of 105 participants, 71 females and 34 males, aged 18–40 ( $M = 20.4$ ,  $SD = 5.01$ ) were tested.

### *Materials*

A set of written material was created for two learning conditions. Participants in the No Causal Knowledge (NCK) 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 (M.G.) and involved three main sections. The first two sections, designed as a control sections, contained an introduction to the anatomy of the lungs and to three diseases involving the lungs (chronic obstructive pulmonary disease, pneumonia and pleural effusion). The third section, created for the testing condition, contained the typical

textbook description of how a physician performs and interprets the three main components of the respiratory physical exam: auscultation of breath sounds, percussion of the thorax and feeling for tactile fremitus.

In addition to what the NCK group had, the Causal Knowledge (CK) group had an extra section in their booklet titled “The Basic Physics of Lung Sounds”. This extra section included the physics explanations for why any given physical finding occurred during each of the components of the respiratory physical exam. For example, all participants read about the exam and how it would sound in a patient with chronic obstructive pulmonary disease (COPD). Participants in both groups were given information about the disease, and the symptoms, and were told what examination findings would be suggestive of COPD. However, (as an example) participants in the CK group were given information like “imagine tapping in the same way on a drum. If it were 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.” In order to balance the overall amount of material presented, the NCK group received an increase in the amount of information provided in the introductory sections of the three lung conditions.

Participants’ understanding and retention was measured via a multiple-choice test that included fifteen test items and ten control items. Test items were questions that required the participant to make a diagnosis based on exam results, or to predict exam findings given a particular condition. As such, we expected that performance on these items would be aided by the causal 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 dependent on the presence of causal knowledge. An example for both the test and control questions is shown in Fig. 1. Note that the surface details of the test item need not explicitly bring to mind the information in the causal knowledge condition (e.g. the

<p><u>Sample Test Item:</u></p> <p>George Martin, a 47 year old male, comes to the hospital by himself complaining of shortness of breath. Upon examination, you discover that over the right lower portion of his lung he has decreased breath sounds, decreased tactile fremitus, and dullness to percussion. Upon reviewing the results you diagnose the patient with:</p> <p>a) COPD  <b>b) Pleural Effusion</b>  c) Pneumonia  d) Normal Lungs</p>
<p><u>Sample Control Item:</u></p> <p>The separation between the two outside layers of the lung is referred to as:</p> <p>a) Alveolar Duct  b) Visceral Pleura  <b>c) The Pleural Space</b>  d) Bronchial Space</p>

**Fig. 1** Sample test and control questions. This figure shows an example of one the test items and one of the control items. The correct answer is shown in bold-face text. When the items were presented to participants, the correct answers were not indicated in any way

fluid-filled drum) but the underlying concepts reflect the causality of the fluid and the sound produced of the percussing. For this item there are three ways to arrive at a correct answer: a guess, memory for the exact combination of findings for pleural effusion, or reliance on the physics principles (e.g., fluid around the lung is dull to percussion and dampens sound transmission). In short, these test items should be sensitive to the causal knowledge (physics) that is presented to the experimental group of participants.

Control items were ones where answering correctly would not be helped by knowing the physics information presented to the CK group. In order to ensure that the immediate and delayed tests were of equivalent level of difficulty, the same test items were repeated in both tests. However, the item's context irrelevant details (e.g. age, gender etc.) and the question sequence were altered in order to make the two tests appear different.

### *Procedure*

Participants were randomly assigned into one of the two learning groups. Participants were tested in small groups of up to six individuals in two sessions held 1 week apart. These sessions took place either in the Categorization Lab at UWO or in the Psychology Research Lab at GMU. During the first session, informed consent was obtained from each participant, and information on participants' English proficiency was also collected to control for possible differences in language ability and understanding. Participants 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 participants could study the booklet or on either test; however, most took between 20 and 30 min to study the material and 10–15 min to do each test. The time it took each participant 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 participant 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 (J.P.M). Sixteen participants were tested with a pencil and paper version of the test: their performances did not differ statistically from the main group.

After completing the test, participants booked their second session for 1 week later. During the second session, participants were given only the computerized test without being presented with the information booklet. After completing the experiment, participants were debriefed. The average time for completing the first session was 42 min and the average time for completing the second session was 10 min.

### *Analysis*

In order to ensure that the developed test items were performing as expected, item- total correlations were performed using a point biserial correlation. Mean scores were calculated for each component (control and test items) of both Test 1 and Test 2. Group differences were then explored using a mixed factorial ANOVA with Condition (CK/NCK) as a between subjects factor and Test (Test 1/Test 2) as a within subjects factor. A t-test was also used to confirm the group differences at Time 1 and Time 2.

**Table 1** Performance on the multiple choice test

	Time 1		Time 2	
	Mean	SD	Mean	SD
Control items				
Causal knowledge	0.76 (.77)	0.11	0.69 (.69)	0.13
No causal knowledge	0.74 (.73)	0.12	0.71 (.70)	0.12
Test items				
Causal knowledge	0.73 (.72)	0.17	0.61 (.61)	0.20
No causal knowledge	0.61 (.60)	0.18	0.49 (.49)	0.22

The values in parentheses are the means from the full sample

## Results

All but four participants indicated that English was their native language. Of those four that were not native speakers, there were two in each experimental group, and the average number of years of English proficiency was 6.8.

Our preliminary analysis included 104 participants who were run in the task (the data from 1 participant was excluded *a priori* because this participants had a university physics background). Due to the nature of using undergraduate psychology students, it is not uncommon that they may sign up for a study but fail to complete it, show a poor effort in completing the task, or guess during the test phase. *A priori*, we decided to exclude participants who did not return, indicated that they mostly guessed, or performed poorly on the immediate test. As such, we restricted our main analyses to a subset of the participants. In all cases, though, the statistical effects were reliable before and after the data restriction, and the means for the full set are shown in Table 1 in parenthesis.<sup>1</sup> The data restriction proceeded as follows. Twenty-three participants were removed from the initial data set for the final analysis. Sixteen participants (11 from the CK group and 5 from NCK) were removed for failing to return for the second test, and 6 participants (1 from the CK group and 5 from the NCK group) were removed because they had an average score of less than 0.40 on the immediate test. One participant was removed from the NCK condition because he had a background in university-level physics. The final analysis consisted of data from 82 participants, 60 females and 22 males.

We examined first the correlation between each participant's overall score and their score on each individual item. The average correlation was 0.34 (a range from 0.02 to 0.54), indicating that our items were adequately diagnostic. The one low item (0.02) was a control item, and so we retained it in our test. We also calculated Cronbach's alpha for the test items as a measure of internal consistency. The score ( $\alpha = 0.70$ ) indicated that the test had an acceptable level of internal consistency.

We scored separately Test 1 and Test 2 and calculated the proportion correct on the set of control items and the set of critical test items for each participant. Performance on the tests is shown in Table 1. Both groups performed well on the control items and there were

<sup>1</sup> The text reports the analysis of the restricted data set. The comparable statistics for the full data sets (drop outs and non learners included) found no effect of condition on the control items,  $F(1, 186) = 0.690$ ,  $p = 0.417$ , and no significant difference between the immediate and delayed control item test scores,  $F(1, 186) = 2.53$ ,  $p = 0.113$ . With regard to the critical test items, a significant main effect of condition was found,  $F(1, 186) = 8.632$ ,  $p = 0.004$ . There was also an effect of Test,  $F(1, 186) = 4.181$ ,  $p = 0.042$ , and no interaction was found  $F(1, 186) = 0.089$ ,  $p = 0.766$ . These statistics differ quantitatively from those reported the main text but do not differ qualitatively.

no significant differences in performance by either group  $F(1, 158) = 0.0172, p = 0.90$ , indicating that (as predicted) the control items were not affected by the addition of causal knowledge. As well, performance on the control items did not decrease with time and there was no significant difference between the immediate and delayed control item test scores,  $F(1, 158) = 2.32, p = 0.13$ .

With regard to the critical test items, a significant main effect of condition (performance was better for the CK group than for the NKC group) was found,  $F(1, 158) = 18.16, p < 0.01$ . There was also an effect of Test (performance on Test 1 was better than on Test 2),  $F(1, 158) = 7.76, p < 0.01$ ; however, no interaction was found  $F(1, 158) = 0.014, p = 0.90$ . Comparison t-tests on the test items indicated a significantly superior performance of the CK group at Time 1,  $t(80) = 2.405, p < 0.01$ , and at Time 2,  $t(80) = 1.625, p < 0.01$ .

## Discussion

The performance seen on both the control and test items suggests that our students learned their materials relatively well. As can be seen from the sample questions in Fig. 1, these were not questions one could do well on by guessing or based on prior experience. Most importantly, the group given the basic science explanation (CK group) had a significantly better performance on test items (but not on control items) during the first time period and the second time period. These findings are in line with our first hypothesis that teaching students, in an integrated fashion, the basics of acoustics as it relates to the respiratory exam will enhance their ability to interpret the findings from a physical exam in various medical conditions.

This finding has two main implications. First, on the simplest level, 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. Second, 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 regard to our study because, unlike previous studies, this is one of the first of its kind to consider physics (acoustics) as a basic science relevant to the teaching of medicine. And, while we are not suggesting that all medical students study physics, we are suggesting that a decision to include certain basic science teachings into the curriculum should take into account their ability to support causal knowledge learning. For example, consistent with our main hypothesis, research by Norman and colleagues has suggested that providing non-standard conceptual examples (mechanical examples in addition to just physiological examples) can improve clinical explanations (Norman et al. 2007).

Importantly, 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 et al. 2010). In the latter study, 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 would have called conceptual coherence (Murphy and Medin 1985).

While our results were generally in line with our hypothesis, we also initially speculated that the CK advantage (the difference between the CK and the NCK groups) might increase over time, as has been shown in other research (Woods et al. 2007, 2005, 2006). That is, we expected that the one-week delay might have a greater deleterious effect on participant in

the NCK condition. However, the CK advantage that we found was strong in both the first and second tests in and we did not find any evidence of an interaction between condition and time. This suggests that although performance did decline after the one-week delay, the conceptual advantage of the causal knowledge manipulation was robust. Additional research in this area is needed to investigate how and why a CK advantage might increase over time, rather than just hold firm.

One possible limitation is that our students had no intrinsic motivation for learning the material, because unlike medical students, the content had no relevance to their future careers. This limitation has some implications for training, because if causal knowledge is to be helpful in acquiring medical knowledge, it must be sufficiently salient and well integrated to allow for more extensive influence. As such, a future study using medical students is warranted.

Another limitation is that we had only minimal control over the timing of the study material and of the test. Although we attempted to equate the amount of text that both groups read, we did not track how long each group spent on average with the training material. It is possible that participants in the CK group simply spent more time with the training material and that extra time, rather than the causal knowledge, produced the better performance by the CK group. We do not favor that interpretation because the better performance by the CK groups was seen only for the critical test items and not for the control items. This result is more in line with the causal hypothesis. We do recommend that future research consider controlling for and examining study time as a potentially important variable.

In conclusion, this study has shown that providing a basic science explanation for clinical phenomena can improve a student's interpretation and recall of clinical information. The eventual implication of this work, and of some of the additional studies we have suggested, could warrant an expansion of which basic sciences are considered basic to medicine and what might make them relevant. In particular, our findings support a recommendation 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 of causal explanations for clinical findings.

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

## References

- Bickley, L., Szilagy, P. G., & Bates, B. (2007). *Bates' guide to physical examination and history taking*. Philadelphia, PA: Lippincott, Williams & Wilkins.
- Finnerty, E. P., Chauvin, S., Bonaminio, G., Andrews, M., Carroll, R. G., & Pangaro, L. N. (2010). Flexner revisited: The role and value of the basic sciences in medical education. *Academic Medicine*, *85*, 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*, 220–227.
- Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227–254.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.
- Norman, G. (2005). Research in clinical reasoning: Past history and current trends. *Medical Education*, *39*, 418–427.
- Norman, G. R., Trott, A. L., Brooks, L. R., & Smith, E. K. M. (1994). Cognitive differences in clinical reasoning related to postgraduate training. *Teaching and Learning in Medicine*, *6*, 114–120.
- Norman, G., Eva, K., Brooks, L. R., & Hamstra, S. (2006). Expertise in medicine and surgery. In K. A. Ericsson, N. Charness, P. L. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 339–353). Cambridge, UK: Cambridge University Press.

- Norman, G. R., Dore, K., Krebs, J., & Neville, A. J. (2007). The power of the plural: Effect of conceptual analogies on successful transfer. *Academic Medicine*, *82*, S16–S18.
- Patel, V. L., & Groen, G. J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, *10*, 91–116.
- Rehder, B. (2010). Causal-based classification: A review. In B. H. Ross (Ed.), *The psychology of learning and motivation*, vol. 52. (pp 29–116). Burlington, VT: Academic Press 52, 29–116.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*, 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*, 611–621.
- The Association of Faculties of Medicine of Canada. (2009). *The future of medical education in Canada (FMED): A collective vision for MD education*. Ottawa, ON: The Association of Faculties of Medicine of Canada.
- 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.
- Weslby, P. D., & Earis, J. E. (2001). Some high pitched thoughts on chest examination. *Postgraduate Medical Journal*, *77*, 617–620.
- 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*, *39*, 107–112.
- 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*, 973–979.
- Woods, 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*, 417–426.