Self-Explanation as a Strategy for Supporting the Development of Diagnostic Reasoning in Medical Students: An Exploratory Study on Knowledge Development

Christian Muhoza-Butoke  
*Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada*

Christina St-Onge  
*Center for Health Sciences Education, Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada*

Martine Chamberland  
*Center for Health Sciences Education, Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada*, martine.chamberland@usherbrooke.ca

Follow this and additional works at: [https://hpe.researchcommons.org/journal](https://hpe.researchcommons.org/journal)

Part of the Health and Physical Education Commons

**Recommended Citation**

DOI: 10.1016/j.hpe.2017.03.005  
Available at: [https://hpe.researchcommons.org/journal/vol4/iss2/3](https://hpe.researchcommons.org/journal/vol4/iss2/3)

This Original Research Reports is brought to you for free and open access by Health Professions Education. It has been accepted for inclusion in Health Professions Education by an authorized editor of Health Professions Education.
Self-Explanation as a Strategy for Supporting the Development of Diagnostic Reasoning in Medical Students: An Exploratory Study on Knowledge Development

Christian Muhoza-Butoke\textsuperscript{a}, Christina St-Onge\textsuperscript{b}, Martine Chamberland\textsuperscript{b,\ast}

\textsuperscript{a}Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada
\textsuperscript{b}Center for Health Sciences Education, Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada

Received 21 December 2016; received in revised form 19 March 2017; accepted 28 March 2017
Available online 19 April 2017

Abstract

Purpose: Self-explanation (SE) has been shown to support the development of diagnostic reasoning in medical students. However, no study has documented how SE influences the development of the knowledge that underpins this reasoning. This study was a first step to explore this. More specifically the aim was to compare the use of biomedical and clinical knowledge by medical students who use SE while solving the same clinical cases one week apart.

Methods: Fifty-three medical clerks previously took part in a study to determine the impact of combining SE with listening to examples of SE on solving clinical cases one week later. In the present study, the authors analyzed the SE verbatim of the 15 students in the control group, who only used SE while solving cases. Four cases per participant were analyzed, for a total of 60 transcribed SE recorded at two different times, one week apart (T1 and T2), and for two clinical cases. The verbatim transcripts were coded according to a pre-determined coding grid: paraphrases, clinical inferences, biomedical inferences, monitoring, and errors. Code frequencies were compared at T1 and T2 using a paired t-test.

Results: No significant difference between the two times in any of the categories: clinical inferences (p=0.28), biomedical inferences (p=0.08), paraphrases (p=0.97), monitoring (p=0.60), and errors (p=0.65).

Discussion: Our results did not show quantitative changes of biomedical or clinical knowledge expressed by students using SE when tested at one week interval. The level of students and the short observation period may explain the negative findings. Alternatively, the assumption that knowledge transformation could be captured by simple quantitative measures might be too simplistic. To document and qualify the effects of SE on the medical knowledge, future studies will need to combine different instruments and/or observe its development over a longer period of time.

Keywords: Clinical reasoning; Diagnostic reasoning; Knowledge building; Self-explanation
1. Introduction

In the past decades, several researchers have focused on the multiple facets of clinical reasoning, a core competency in medical practice. One key finding is the core role of specific knowledge in clinical reasoning. Clinical reasoning is developed gradually over the course of a student's medical training, with repeated clinical exposure to actual patients playing a central role in this development. Medical students' transition from the preclinical phase of their training to their clerkship is a potential turning point for partial integration of the knowledge needed to develop clinical reasoning. The challenge for programs and teachers is to seize these opportunities and to implement strategies that will enable students to actively build their knowledge based on real clinical situations that they have experienced.

Schmidt's theory of expertise describes the development and restructuring of students' knowledge over the course of their training. According to this theory, students go through four transitional stages in the development of their expertise in medicine. The first stage is characterized by knowledge in the form of rich, elaborate causal networks that explain the causes and consequences of disease in terms of underlying biological or pathophysiological processes. The second stage occurs when the students are exposed to clinical problems. This causes their networks of causal knowledge to gradually become "packaged" under more abstract concepts in a phenomenon known as "knowledge encapsulation." The third stage is characterized by the emergence and enrichment of illness scripts, cognitive entities preferentially containing clinical knowledge about the characteristics of disease, such as predisposing factors, common clinical manifestations (signs, symptoms), and a brief description of the dysfunction. At this stage, the "encapsulated" biomedical knowledge is not strongly expressed but contributes to the coherence of the script. At the fourth stage, the knowledge acquired during the previous stages is enriched by the addition of clinical experiences with actual patients.

One of the techniques studied in various fields and put forth to support learning when solving problems is self-explanation (SE). SE is an active learning technique that consists in the student generating explanations for himself based on learning material (e.g., a text, a problem, an example of a solution to a problem) in order to improve his understanding. SE enables the student to develop his knowledge, generate inferences, integrate new information to his previous knowledge, monitor his knowledge and, consequently, review its representations stored in memory. Also, in a recent literature review in the field of science and mathematics, Richey & Nokes-Malach examined how different techniques, including self-explanation, specifically contribute to the development of characteristics of expert knowledge. These authors concluded that in these fields, there are data supporting that SE helps to create links between the different types of knowledge, supports the coherence of these links, and helps to identify and resolve contradictions in prior knowledge.

In medical education, SE appears to be a promising learning technique to support the development of diagnostic reasoning in students. In fact, diagnostic performance has been shown to improve in medical clerks who use SE while solving clinical cases compared to those who solve the same cases in silence. This improvement, measured one week later, is observed when the clinical topic approached by the student is less familiar to him, and is more evident when performance is measured for different clinical cases or transfer cases.

Based on the underlying mechanisms of SE and in transposing them to the development of clinical reasoning, this technique could potentially help students to organize their knowledge, make links between their biomedical and clinical knowledge, and review or refine their illness scripts. In fact, when the student generates SE while solving a clinical case that is less familiar to him, he reactivates and expresses his biomedical knowledge more than in a situation of familiarity. It is therefore plausible to think that SE, in a less familiar situation, allows the student to give meaning to the clinical information by making additional links between his biomedical and clinical knowledge, thus allowing him to increase the coherence of the illness scripts that he is in the process of forming. Once the student made connections between the clinical elements and the underlying relevant biomedical knowledge, he or she would not necessarily need to explicitly revisit the later when subsequently facing a similar clinical situation. Similarly, following a narrative review of the teaching techniques for clinical reasoning at the undergraduate level, Schmidt and Mamede put forth the assumption that self-explanation while solving problems could speed up the encapsulation of students' biomedical knowledge.

Despite the evidence of the positive impact of SE on students' diagnostic performance, its effects on specific knowledge in medicine remain hypothetical for the
moment. No study to date has more directly explored the effect of this strategy on the development and/or organization of students’ medical knowledge.

One possible way of observing students’ knowledge in problem-solving situations would be to examine the content of the verbatim they generate during SE. In fact, in analyzing the content of SE, different types of statements become apparent, the impact of which on learning is variable.\textsuperscript{6,7} Statements identified as paraphrases show that the student adheres very closely to the text and have little impact on learning. On the contrary, statements that contain inferences, elaborations that go beyond the content of the text, have a greater influence on learning. The analysis of the verbatim also reveals statements that attest to the student’s monitoring of his ongoing learning. Finally, certain statements may be incorrect but still result in learning if the student realizes his mistake through monitoring and attempts to correct it with self-explanations.\textsuperscript{6} Analysis of SE verbatim has been used in a previous study of medical students to document these different types of statements and, in particular, to describe the type of knowledge—clinical or biomedical—contained in the inferences expressed by the students while solving problems on topics that are familiar and unfamiliar.\textsuperscript{12} It is therefore plausible that we might observe, at least partially, the development of students’ knowledge through the analysis of their SE verbatim.

In this study, we explored the development of knowledge in medical students who used SE in problem-solving situations by analyzing the content of their SE verbatim generated on the same cases one week apart. We put forth the assumption that if self-explanation facilitates the encapsulation of biomedical knowledge, the students would express less biomedical knowledge one week later relative to the learning phase.

2. Method

2.1. Design

In this exploratory study, we analysed the content of SE verbatim generated by students while solving clinical cases one week apart and compared the types of knowledge they verbalize with a particular interest on biomedical and clinical knowledge.

2.2. Context/Participants

The current study was done using SE verbatim generated by students and collected during a previous study.\textsuperscript{13} This experimental study of 53 students measured the effect of three conditions on the students’ diagnostic performance: the use of SE followed by listening to a peer's SE, the use of SE followed by listening to an expert's SE, and the use of SE alone. As the purpose of this study is to explore the development of knowledge by students who use SE alone, our analysis will focus exclusively on this latter group of students.

The participants were students enrolled in the medicine program at Université de Sherbrooke, in Sherbrooke, Québec, Canada. This program involves 2.5 years of problem-based learning (PBL) and an 18-month clerkship. At the time of the study, the students were in the third year of their training, more specifically, at the clerkship in an internal medicine rotation. Participation in the study was voluntary and the students signed a consent form. Université de Sherbrooke’s Ethics and Research Education, and Social Sciences Committee approved the original study,\textsuperscript{13} and renewed its approval for the current study.

The original study consisted of two phases: a learning phase and an assessment phase one week later. In the learning phase, after a brief training on SE, the participants in the SE alone group solved four separate clinical cases aloud. In the assessment phase a week later, the participants solved the same four clinical cases as during the learning phase, again with verbal SE and eight different cases in silence. At the end of each clinical case, the students were asked to write down their main diagnosis, two arguments supporting the diagnosis, and two plausible alternative diagnoses. For all participants in both phases, the audio of their SE was recorded. The group of students using SE alone included sixteen participants. Among those, one student was excluded from the analyses because the recording of his verbatim was incomplete. Fifteen students whose recordings were complete were selected for this study.

2.3. Material

The four clinical cases for which the students used SE in making their diagnoses in both phases were all related to jaundice, but involved different pathophysiological mechanisms and disease categories, i.e., acute hepatitis, chronic liver disease, obstructive jaundice, and hemolytic jaundice. Jaundice was considered a priori as a less familiar clinical topic for the participants, given their prior exposure in their medical curriculum. These clinical cases are described in detail in the original study.\textsuperscript{13} Given the large volume of
Clinical case 1
76-year-old man consulting for progressive jaundice

History of current illness: Patient notes the gradual appearance of a greenish-yellow tinge to his skin over the past 3 weeks. Aside from vague, post-prandial epigastric discomfort, he has no abdominal pain. He has noticed that his urine is darker and his stools lighter, with no associated diarrhea or constipation. His appetite seems to have decreased and he says he's lost about 5–7 kg in the past month. He has no fever, chills, or night sweats. No travel or change in eating habits.

Past medical history: Systolic hypertension and type 2 diabetes for 1 year. Acute diverticulitis in 2006. Superficial thrombophlebitis in right leg two months ago. Medications: hydrochlorothiazide 12.5 mg q.d, metformin 250 mg b.i.d.

Lifestyle: Retired carpenter, still lives at home, widower x 3 years. He does not smoke or drink alcohol.

Physical examination: Patient in good general condition, BP 145/75, pulse 75/min., afebrile, weight 75 kg. Obvious jaundice of the skin. No lymphadenopathy. Examination of the heart and lungs is normal. No painful area on palpation of the abdomen. The liver and spleen appear to be normal in size, with no palpable mass. Rectal examination is normal. No edema in the legs.

Paraclinical work-up: CBC: hemoglobin 125 (130–180 g/L), MCV 88 (80–100), WBC 6.2 (3.8–10.6 × 10^9/L), platelets 350 (130–400 × 10^9/L); electrolytes and creatinine are normal; fasting glucose 6.9 (3.3–6.1 mmol/L); AST 52 (14–50 IU/L), ALT 75 (21–72 IU/L), alkaline phosphatase 625 (43–200 IU/L), total bilirubin 110 (3.4–17 umol/L), conjugated 100; albumin 33 (35–50 g/L); INR and aPTT normal. Urinalysis shows brown urine, highly positive for bilirubin and negative for urobilinogen.

Clinical case 2
50-year-old patient consulting for jaundice and fatigue

History of current illness: Patient complaining of progressive fatigue for the past two months and a yellow tinge to his skin for the past week. No abdominal pain, appetite and weight are stable. No specific digestive symptoms, stools and urine are normal in colour. No fever, chills, or sweating, but complaining of dyspnea on effort for the past two weeks with no cough or chest pain.

Past medical history: Appendectomy as a child, Hashimoto’s thyroiditis with secondary hypothyroidism for the past 10 years. No specific family history. Medications: levothyroxine 0.1 mg q.d, acemetacin phen pm, no over-the-counter medications.

Lifestyle: Patient has smoked a pack a day for 30 years, no drugs or alcohol, no high-risk sexual behaviour, construction worker, no recent travel.

Physical examination: BP 130/80, pulse 105/min., temperature 37.8; eupneic at rest; O2 saturation 98% on room air. Jaundice of the skin and sclera. Pale skin. No lymphadenopathy. Examination of the heart shows a systolic murmur 2/6 in left parasternal, non-radiating, with normal sounds. Pulmonary auscultation is normal. The abdomen is soft with no hepatomegaly or mass. Spleen edge is palpable on deep inspiration. No peripheral edema.

Paraclinical work-up: CBC: hemoglobin 62 (130–180 g/L), MCV 102 (80–100), WBC 11 (3.6–10.6 × 10^9/L), platelets 400 (130–400 × 10^9/L); creatinine, electrolytes, and glucose normal; AST 48 (14–50 IU/L), ALT 60 (21–72 IU/L), alkaline phosphatase 198 (43–200 IU/L), total bilirubin 68 (3.4–17 μmol/L), conjugated 5; INR and aPTT normal; urinalysis negative for bilirubin and highly positive for urobilinogen.

We had 60 SE verbatim for analysis (15 students × 2 clinical cases × 2 phases) that had already been denormalised at the time of collection. The audio recordings of the SE verbatim were transcribed by a team of independent professionals. In the text that follows, we will use the term “SE protocol” to describe the document that corresponds to the verbatim transcribed for each clinical case.

The SE protocols were then randomly sorted into a list from 1 to 30 for each of the two clinical cases, to ensure a blind analysis of the protocols by the evaluators, meaning the latter would have no way of knowing to what phase (learning or assessment) or what student the protocols corresponded.

To analyze the SE protocols, we used the coding grid developed previously and inspired by the technique described by Chi. For each protocol, the sections representing a straight rereading of the text of the clinical case were excluded from the analysis. The rest of the verbatim was segmented into idea units. One idea unit corresponds to one segment expressing a single idea. Each segment was coded according to the following five mutually exclusive categories: paraphrase, biomedical inference, clinical inference, monitoring, or error. Table 2 presents examples of segments in each category.
The categories are defined as follows:

- **Paraphrase:** A segment was coded as a paraphrase when the participant repeated a section of the original text in his own words without adding any new information. The use of simple medical vocabulary (e.g., dyspnea for shortness of breath) was also considered as paraphrasing.

- **Inference:** A segment was coded as an inference when the participant elaborated, made a link between two ideas, or added information beyond what was explicitly stated in the text.

- **Clinical inference:** An inference related to characteristics of the patient or the disease, to the different clinical manifestations (signs and symptoms), to tests or treatments, or to possible diagnoses.

- **Biomedical inference:** An inference related to a basic mechanism, a physiological function, or a pathophysiological element of an organ or a system.

- **Monitoring:** A segment was coded as monitoring when the participant expressed understanding, lack of understanding, confusion, or doubt about the text, or acknowledged the limits of his current knowledge.

- **Error:** A segment was coded as an error when the participant expressed an idea that was completely incorrect. A segment whose content was ambiguous or partially incorrect was not coded as an error.

Two evaluators [CMB (senior internal medicine resident) and MC (internal medicine specialist)] coded the SE protocols using NVivo 10 (QRS International Inc., Burlington, MA, USA), a software used to facilitate qualitative data management and analysis. One of the evaluators (CMB) coded all of the protocols.
and the second (MC) coded 20% of the protocols independently. Inter-rater reliability was checked at four moments during coding. First, after coding three protocols for each of the two cases. Disagreements were discussed and all data were recoded. Second, after the 12th, 21st and 30th protocols for each of the two cases were double-coded. Inter-reliability coefficients ranged from 88% to 100%, with only three instances out of 30 being lower than our .9 cutoff. Those instances were discussed and yielded additional refinement of the coding grid. The revised coding grid was then applied to all the data set by one author (CMB).

For each protocol, the number of coded segments in each category was automatically totalled by NVivo. Once the coding was finalized, a research assistant re-matched, for each participant and each case, the protocols from the learning phase and the assessment phase. Paired t-tests were performed on the number of segments in each category, for both cases. The significance threshold was set at 0.01 after the Bonferroni correction.

3. Results

Table 3 presents the mean number of segments in each category for both clinical cases combined. The paired t-tests show no significant difference for each of the categories analyzed.

4. Discussion

The aim of this study was to explore the effect of self-explanation (SE) on knowledge-building during diagnostic reasoning in medical students, with the specific objective of comparing the use of clinical and biomedical knowledge by a student who is using SE while solving the same clinical case one week apart. Based on the assumption that SE facilitates knowledge encapsulation, we expected that the student would, for the same case, demonstrate less biomedical knowledge a week later. However, the results of our comparative analysis of the content of SE protocols showed no significant difference between the mean numbers of statements in each of the coding categories for the two measurement points.

Our results must not be interpreted as a demonstration of the absence of SE effect on the types of knowledge expressed by students. Non-negligible variance was observed among the students, as shown by the large standard deviations associated with the mean numbers of segment categories. This variance makes it less likely to observe significant differences between the two measurement points. Segmentation and coding were systematically double-checked to ensure that there was no major error in coding that could explain this high variance (for example double coding). Since the participants were all in the same cohort, enrolled in the same clinical rotation at the time of the study, and had the same clerkship rotations schedule, it is also unlikely that the variability stems from the students’ exposure to clinical problems. The participants in our study were initially introduced to the technique in a standardized manner, but they were not given any additional specific instructions about self-explanation during the sessions. Therefore, the observed variance between participants might reflect the individual variability of self-explanation. In fact, the quantity of SE statements generated by participants spontaneously or even when prompted has been reported to vary significantly. This factor may have contributed to the low statistical power observed in our study. Analyzing all 120 transcriptions generated for the four cases available would not have increased the statistical power as it would not have added to our pool of participants.

Another explanation for our negative results might be that a quantitative change in the type of knowledge expressed by the student may not be observable after only one week or after only two SE sessions for a specific clinical case. Otherwise, while we assumed that knowledge encapsulation or enhanced coherence of mental representation of diseases would translate into different amount of ideas, maybe it is not the case and we could have rather looked more for qualitative differences. While our coding grid used to categorize SE segments allowed us to quantify biomedical and clinical inferences in time, it did not allow for a detailed qualitative analysis of the content of the protocols. Therefore, our instrument may not have been the most appropriate one to detect the changes in students’ knowledge.

We may also consider the specific way SE was operationalized in this study and hypothesize how it might have influenced the results. According to Wylie & Chi, the different forms of SE may be viewed along a continuum depending of how focused is the SE prompt provided to students. In the present study, we used an open-ended SE prompt, which allowed the student to make connections between his prior knowledge (being biomedical or clinical) and any element of the problem. There was no explicit expectation about the preferential expression by students of biomedical knowledge nor constraint to systematically refer to underlying mechanisms while self-explaining. This might explain the finding that over three quarters of the SE segments expressed by students were categorized as clinical inferences. Hence, the open-ended prompt not focussing on biomedical knowledge, might have limited our capacity to observe a difference for
this type of inference over time. Alternatively, the level of the learners could also play a role. Schmidt & Mamede\textsuperscript{4} proposed that SE would speed up knowledge encapsulation but for novice learners. In the present study, the students were at clerkship level and we could assume, based on the theory of expertise\textsuperscript{3} that they had already, at least partially, encapsulated relevant biomedical knowledge and were actively working with illness scripts. In this context, SE could promote knowledge elaboration and making connections between different pieces of clinical knowledge as well as between biomedical and clinical knowledge resulting in no quantitative difference in types of inferences between the two measurements.

Our study was intended as a preliminary exploration of the effect of SE on knowledge building in medical students. We chose the SE protocols as the basic data to be analyzed based on the assumption that the content would be representative enough of the knowledge in development to observe the latter's evolution. By choosing to compare the same student's SE about the same case at a one-week interval, we believed that we had controlled some key variables that would allow documenting this evolution. While the SE protocols generated by the students certainly allow for some insight into the knowledge that is activated and used, they likely only provide partial access to the knowledge in development. Indeed, it is a well-known fact that clinical reasoning is a complex process, and that describing the knowledge upon which it is based is a methodological challenge in itself.\textsuperscript{1,2} In their literature review on SE and other instructional techniques in the field of science and mathematics, Richey & Nokes-Malach\textsuperscript{8} made the point that even if the instructional event for instance SE, is observable, the effect on knowledge can only be inferred. They proposed to make use of a variety of types of assessment in order to develop a more complete understanding of the ways SE as well as other techniques impact on knowledge building.

There is a need to better understand the impact of SE on students’ knowledge building in order for SE to be used effectively and at the right moment in the undergraduate curriculum.\textsuperscript{3} Further studies are required and should explore knowledge using multiple assessment methods simultaneously and over a longer period of time in order to better understand the ways SE impacts on knowledge building.

\textbf{Disclosure}

None.

\textbf{Ethical approval}

Ethical approval has been granted from the Université de Sherbrooke's Ethics and Research - Education, and Social Sciences Committee (CER-ESS 2010-75).

\textbf{Funding}

This work was completed, in part, due to a grant from the Société des médecins de l’Université de Sherbrooke Fund (SMUS-2014-2) for medical education research.

\textbf{Other disclosure}

None.

\textbf{Acknowledgments}

The authors are thankful to Kathleen Ouellet M.Sc. (Center for Health Sciences Education, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada) for her help and guidance during the coding, and with the use of the software.

\textbf{References}


**Christian Muhoza-Butoke** is a senior resident in Internal medicine, Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada

**Christina St-Onge** is associate professor at the Department of Medicine, and scientist at the Center for Health Sciences Education, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada

**Martine Chamberland** is professor at the Department of Medicine, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Canada