Effects of deliberate reflection on students' engagement in learning and learning outcomes

Ligia Maria Cayres Ribeiro,[1] iD Silvia Mamede,[2] iD Eliza Maria de Brito,[1] Alexandre Sampaio Moura,[1] Rosa Malena Delbone de Faria[3,4] & Henk G Schmidt[2] iD

CONTEXT Reflection in practice is assumed to enhance interest in knowing more about a topic, increasing engagement in learning and learning outcomes. However, this claim lacks empirical evidence, particularly in medical education. The authors investigated the effects of deliberate reflection upon clinical cases on medical students' engagement in a learning activity and learning outcomes.

METHODS A three-task (diagnostic task; learning activity; test) experiment was conducted in August 2017. Seventy-two fourth-year students from UNIFENAS-BH Medical School, Brazil, diagnosed two clinical cases with jaundice as the chief complaint, either by following a deliberate reflection procedure or making differential diagnosis. Subsequently, all participants received the same study material on the diagnosis of jaundice. Finally, they took a recall test on the study material. Outcome measurements were study time and test scores.

RESULTS There was a significant effect of experimental condition on students' engagement in the learning activity and on learning outcomes. Students who deliberately reflected upon the cases invested more time in studying the material than those who made a differential diagnosis (respectively, mean = 254.97, standard deviation = 115.45 versus mean = 194.96, standard deviation = 111.68; p = 0.02; d = 0.53). Deliberate reflection was also related to higher scores in the test relative to differential diagnosis (respectively, mean = 22.08, standard deviation = 14.94 versus mean = 15.75, standard deviation = 9.24; p = 0.03; d = 0.51). Medium effect sizes (Cohen's d) were observed in both measurements.

CONCLUSIONS Relative to making differential diagnosis, deliberate reflection while diagnosing cases fostered medical students' engagement in learning and increased learning outcomes. Teachers can employ this relatively easy procedure, possibly both with simulated and real scenarios, to motivate their students and help them expand their knowledge, an important requirement for their professional development.

[1]Department of Medical Education Development, José do Rosário Vellano University (UNIFENAS) Medical School, Belo Horizonte, Brazil
[2]Institute of Medical Education Research Rotterdam, Erasmus MC University Medical Center Rotterdam, Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, the Netherlands

[3]Department of Clinical Pathology and Medical Education, Minas Gerais Federal University (UFMG), Belo Horizonte, Brazil
[4]José do Rosário Vellano University (UNIFENAS) Medical School, Belo Horizonte, Brazil

Correspondence: Ligia Maria Cayres Ribeiro, UNIFENAS Medical School, Libano Street, 66, zip code 31.710.030, Belo Horizonte, Brazil. E-mail: ligiacayres@gmail.com

INTRODUCTION

A hiker in an unfamiliar field faces a branching. Unsure of what direction to follow, he or she scrutinises the options and tries to find evidence in favour of one or another direction, searching references in memory, exploring sights and sounds, using a compass or even climbing a tree. John Dewey[1] uses this simple example to illustrate reflection: the attentive thought we engage in when facing hesitation or doubt; and the deliberate reasoning process of exploring a challenge, from different perspectives, until finding a proper solution for it.[1,2] The hiker has a better chance to find the right way through reflection than by randomly choosing a direction, but only if he or she has relevant knowledge of the context. A compass, for example, will be useless if he or she cannot use it. If this is the case, once (hopefully!) home, he or she might then feel motivated to engage in learning activities to acquire this skill, actually engage in learning and, ultimately, master it. This example can be generalised to other contexts, such as medical education. Because of its potential to trigger engagement in learning, reflection, therefore, can be – and indeed has been – considered an important educational strategy.[3,4] It should be noted, however, that although based on solid theory, there is little empirical evidence that this actually happens, especially in medical education. This article reports on a study that investigated the effects of deliberate reflection while solving clinical cases on medical students' engagement in a follow-up learning activity and learning outcomes.

'Forked-road' situations are common in medical practice because cases are often ambiguous, with patients frequently presenting clinical findings that raise diagnostic uncertainty. In these situations, clinicians may engage in reflection in the midst of the encounter with the patient, while it is still possible to respond to the situation at hand, what Donald Schön called 'reflection-in-action'.[2] For a medical doctor facing a diagnostic challenge, this would involve looking at the patient's problem with different potential explanations in mind, confronting different hypotheses with available clinical data and testing them until setting a diagnostic solution for the patient.[5,6] It is likely that, in the course of this reflective reasoning, the clinicians, especially when they are novice ones,

recognise gaps in their knowledge or just realise that they are not so certain about their choices. This might raise interest in knowing more about the problem, which would consequently trigger engagement in learning activities, resulting in increased knowledge about the topic. Indeed, in a previous experiment, deliberate reflection while solving clinical cases by arguing for and against one's diagnostic hypotheses, listing patient's findings that match and refute them, fostered medical students' situational interest in the cases compared with the more conventional approach of giving differential diagnoses.[7] It is reasonable to assume that this deliberate reflection procedure would foster medical students' learning, because situational interest, a transient and context-related interest also described as 'thirst' for knowledge, has proved to be a good predictor of engagement in learning and learning outcomes in experiments with audiences outside medical education.[8–11]

This would be a promising educational effect of deliberate reflection because engagement in learning activities, whether measured as years of schooling, days of instruction or hours of classes, has repeatedly been shown to improve learning outcomes.[12] Rather than undertaking long-term activities, a medical student or a physician facing a diagnostic challenge is likely to engage, possibly individually, in short-term learning activities such as reading a text on a selected medical topic. However short, this time may make a difference. Research has shown a positive effect on learning outcomes of engagement, even in short learning activities measured in minutes of study time. For example, an experiment on student teachers learning of journal's quality assessment by either inventing or studying a worked solution, found study time to be positively correlated with students' outcomes, even after controlling for prior knowledge and independently of learning strategy.[13] Similar results have been found with secondary school students,[14] and the positive relationship between study time and learning outcomes is reinforced by research showing that reducing students' time to master a new topic decreased their scores on immediate and late tests.[15] Therefore, if deliberate reflection while practising the diagnosis of clinical cases indeed fosters engagement in learning activities and learning outcomes, it would be a helpful tool for the development of medical students' clinical knowledge. To our knowledge, however, there was no empirical evidence that this actually happens.

To address this question, we conducted an experiment with fourth-year medical students, who solved clinical cases by either deliberately reflecting upon them or giving a differential diagnosis. Although some degree of reflection is expected while students think of alternative diagnoses for a clinical case, the deliberate reflection procedure used in this experiment demanded comparing and contrasting clinical data with different diagnoses, and searching for evidence supporting and refuting each one in a systematic way. This structured process is expected to raise more uncertainty than the more conventional approach of generating alternative diagnoses. Subsequently, an appropriate learning task and a test were administered. Engagement in the learning activity, measured as study time, and learning outcomes, measured as scores on a test, were obtained. We hypothesised that: (i) deliberate reflection would foster engagement in learning; and (ii) deliberate reflection would foster learning, relative to providing a differential diagnosis.

METHODS

Design

The study was an experiment with random allocation of participants to either an experimental or a control condition. The experiment consisted of a single session divided into three parts: a diagnostic task, a study task and a recall task. In the diagnostic task, participants diagnosed two clinical cases by following either a structured procedure to reflect upon the cases (experimental) or to make a differential diagnosis (control). After diagnosing the cases, participants from both conditions were presented with the same study material about the differential diagnosis of jaundice. Subsequently, they performed a cued-recall task about the material that they had just studied.

Setting and participants

All 123 fourth-year medical students at José do Rosário Vellano University (UNIFENAS) in Belo Horizonte, Brazil, were invited to voluntarily participate in the study. UNIFENAS has a 6-year problem-based curriculum, with the two final years dedicated to clerkships. We selected fourth year students because at this point in their training they have been exposed to knowledge about patients with jaundice during tutorial groups and lectures but have limited clinical experience with them.

The students who volunteered for the study were recruited as participants and gave written consent. The experiment was run as an extra-class activity in six sessions offered on different days to accommodate the students' timetable. A lottery ticket with an electronic tablet as a prize was offered to the students as an acknowledgement of their participation.

Materials and procedure

A computer-based exercise, consisting of two to-be-diagnosed clinical cases, study material and a cued-recall task, was created for this study using Qualtrics.

The cases had jaundice as the main clinical finding and consisted of a written description of clinical symptoms, physical examination and laboratory test findings. Each case had a most likely diagnosis that had been validated in previous studies.[16,17] The diagnoses of the cases were acute viral hepatitis and choledocholithiasis.

The study material consisted of an illustrated text presenting a brief review of bilirubin physiology and physiopathology, followed by the presentation of the clinical cases participants had diagnosed, with the key clinical findings valuable for differentiating between the causes of jaundice highlighted and linked to boxes with their interpretation and explanation (e.g. 'Coluria is an indicator of cholestasis because only free direct bilirubin is excreted in urine').

To measure students' learning of the study material, a cued-recall task with eight open-ended questions, addressing topics on clinical history, physical examination and laboratory tests concerning the differential diagnosis of jaundice, was created. Each item cued recall of a specific part of the material and, for that part, requested the participants to write down all the information that they remembered from the text that they had just read. An example of a cued question is 'List all the relevant physical examination findings to the evaluation of patients with jaundice and explain how they help on the differential diagnosis'. A pilot of the whole exercise was previously conducted with 15 fourth-year medical students, non-participants in this study, to check for understandability, and resulted in a few slight adjustments.

At the end of the exercise participants were asked to answer questions on gender, age and an

estimation of their previous experience with jaundice, measured by how many real patients with jaundice they recalled having assisted.

In both conditions, the exercise was presented to students on a computer screen in the following sequence: (i) a clinical case followed by different instructions to work on the case, depending on the experimental condition to which they were assigned (see below); (ii) the study material; (iii) the cued-recall test; and (iv) the personal data and jaundice experience questionnaire.

For each case, the students were asked to follow a set of different steps depending on the experimental condition under which they performed. In the deliberate reflection condition, students were requested to: (i) read the case and type down the most likely diagnosis for the case; (ii) reflect upon the case by following a structured procedure,[6] which, briefly, consists of typing the clinical findings that are in line with their first diagnostic hypothesis, those that contradict it, and those that are expected were this first hypothesis true but are not described in the case, then typing two alternative diagnoses and running the same analysis for each diagnosis before making a conclusion. Students from the control condition were asked to carry out the following sequence of steps: (i) read the case, type down the most likely diagnosis for the case and two alternative diagnoses; (ii) work on a crossword containing medical terms not related to the cases; and (iii) type down their conclusion on the most likely diagnosis for the case. The crossword was used to ensure a similar time on task and to minimise the degree to which those in the control condition would engage in reflection. As working on a crossword after diagnosing a case is unusual for students, which could hinder compliance, we stated in the instructions to the control condition that, although it could seem irrelevant, a task like it might help the thinking process and should, therefore, be taken seriously. The instructions for each experimental condition were presented exclusively on screen, thereby preventing crossing instructions between students who performed under one or the other condition.

The experiment was conducted in a computer laboratory with partitioned seats keeping students working individually. Two different electronic addresses were created in Qualtrics, each one forwarding to the experimental or the control condition exercise. After a brief introduction to the study, printed electronic addresses, which had been previously organised in random order, were distributed to the students, thereby randomly assigning participants to one of the two conditions (reflection or control). The students worked on the same exercise throughout the experiment, with each student performing therefore only under the condition to which he or she was initially assigned. This procedure for randomisation was chosen because it was not possible to anticipate which students would volunteer for the study and actually attend the activity. A teacher was present in all sessions both to support the participants and to inhibit students from consulting resources other than the exercise.

Time and progress throughout the sequence of steps were computer controlled. After the instructions, time was allocated as follows: 2 minutes to type the most likely diagnosis (reflection condition) or to type the most likely diagnosis and two alternative diagnoses (control condition); and at least 4 to 6 minutes to reflect upon the case (reflection condition) or to solve the crossword (control condition). There was no fixed time to work on the study material or on the cued-recall task. The software automatically recorded time spent on each task and responses for each participant. The participants could neither skip steps nor move backward on the exercise but were free to progress at their own pace while studying the material and carrying out the recall task. They did not receive any information about their diagnostic performance while working on the cases. The correct diagnoses were explained in the study material.

Data analysis

The diagnoses provided by the students in the diagnostic task were firstly independently assessed by two board-certified internists (LMCR; EMB), who were not aware of the experimental condition under which they had been made. They classified each response as: correct (scored 1), whenever the core diagnosis was present (e.g. 'hepatitis' in the 'acute viral hepatitis' case); partially correct (scored 0.5), if the core diagnosis was not present but a component of it was (e.g. 'gallstones' in the 'choledocholithiasis' case); and incorrect (scored 0), when the response did not fall into any of these categories. The raters agreed on the score attributed in 87% of the responses and resolved discrepancies by discussing them and reaching consensus in a subsequent meeting. The accuracy of initial diagnoses was measured to check whether the

two experimental groups were similarly acquainted with the problem under study.

The amount of information reported in the participants' responses (recall task) was evaluated by counting the number of idea units present in each response.[18,19] We considered as idea units, text fragments, such as a word or a short sentence, meaningful to the task. Because we were interested in measuring learning of the information presented in the study material and used a cued-recall task, the idea units counted in participants' responses had to meet three criteria: to consist of correct information, to be actually present in the study material and to be consistent with the question cue. Each idea unit that met these criteria was scored with one point. Two authors (LMCR, EMB) firstly discussed and then reached a consensus on the idea units present in the study material. One example of an idea unit, consistent with the cue 'Explain how urine and feces' color help differentiating the causes of jaundice' is 'coluria indicates cholestasis'. Subsequently, the same authors independently assessed 10% of participants' tasks, reaching an initial agreement of 86%. After discrepancies were resolved in further discussions, the first author assessed all participants responses. Only manifest content was considered: no possible underlying meanings were counted.[20]

The main outcome measurements of the study were the time students spent on the study material, which has been taken as a measure of the learning process, and their scores on the cued-recall task, assumed to be a measure of learning outcomes. The mean study time was computed for each experimental condition. The scores on the cued-recall task were computed by summing the scores of each participant and, subsequently, averaging them for each experimental condition. Mean diagnostic accuracy scores were computed through a similar procedure. Two separate ANOVAs with experimental condition (reflection or control) as between-subjects factor were performed on the mean study time and on the mean cued-recall task scores; *t*-tests were performed to compare age and diagnostic accuracy and Pearson chi-squared tests to compare gender and previous experience with real patients with jaundice between experimental and control groups.

RESULTS

Seventy-four (60%) out of the 123 invited students participated in the study. One outlier was removed from the control group after the exploratory data analysis. One participant was removed from the reflection group for accessing text references other than the experiment material during the exercise, which was not allowed. This led to 72 participants, 36 in each group condition.

Table 1 presents age, gender, previous experience with patients with jaundice and initial diagnostic accuracy for the clinical cases as a function of experimental condition. No significant differences emerged in age, $t(54.11) = 1.35$, p = 0.18, or gender, $\chi^2(1) = 0.53$, p = 0.63. Previous experience with patients with jaundice, $\chi^2(1) = 0$, p > 0.99, and the initial diagnostic accuracy scores for the clinical cases, $t(70) = 0.50$, p = 0.61, also did not differ between groups, indicating that they were comparable.

The mean study time and cued-recall task scores for the two experimental conditions are displayed in Table 2. There was a significant main effect of experimental condition on study time, $F(1, 70) = 5.03$, p = 0.02, $d = 0.53$, with students who deliberately reflected upon the cases engaging longer in the subsequent learning activity than those who had given differential diagnoses. There was also a significant main effect of experimental condition on learning outcome, $F(1, 70) = 4.68$, p = 0.03, $d = 0.51$, with students who deliberately reflected upon the cases showing higher scores on the cued-recall task than those who gave differential

Table 1 Characteristics of the participants in the experimental and control groups (standard deviation in brackets)

	Reflection group	Control group	Overall
Age (mean)	22.81 (1.83)	23.67 (3.36)	23.24 (2.72)
Gender			
Male	12 (33%)	15 (42%)	27
Female	24 (77%)	21 (68%)	45
Previous experience with patients with jaundice			
Yes	21 (58%)	21 (58%)	42
No	15 (42%)	15 (42%)	30
Initial diagnostic accuracy scores on clinical cases (range 0–2)	1.29 (0.74)	1.37 (0.66)	1.33 (0.69)

Table 2   Mean study-time (seconds) in the learning activity and cued-recall scores in the test (total number of information units) as a function of experimental condition (standard deviation in brackets)

	Reflection group	Control group	Overall
Time spent in the learning activity	254.97 (115.45)	194.96 (111.68)	224.96 (116.76)*
Cued-recall score obtained in the test	22.08 (14.94)	15.75 (9.24)	18.92 (12.74)†

* Significant main effect of experimental condition (p = 0.02, *d* = 0.53).
† Significant main effect of experimental condition (p = 0.03, *d* = 0.51).

diagnoses. The effect sizes, as measured by Cohen's *d*, were medium for both outcomes (considering *d* values of 0.2 for small, 0.5 for medium and 0.8 for large effect sizes).[21]

DISCUSSION

In this study, we investigated the effect of deliberate reflection while solving clinical cases on medical students' engagement in a learning activity and learning outcomes. To that end, we compared study time and test scores of students who worked with two clinical cases through deliberate reflection with those who made differential diagnoses. Students who worked with deliberate reflection engaged 31% longer in the learning activity and attained a 40% higher score on the cued-recall test. The difference in favour of the deliberate reflection condition was significant in both measurements, with a medium effect size.

These results are in line with our expectations: students who deliberately reflected upon to-be-diagnosed clinical cases had a more challenging task than those who worked with differential diagnosis. It is reasonable to expect that students who are requested to provide alternative diagnoses for a case engage in some degree of reflection to perform the task. The deliberate reflection, however, requires students to compare and contrast the alternative diagnoses, searching for evidence supporting and refuting each one in a systematic way. This process might raise uncertainty and facilitate recognition of knowledge gaps, which has already been shown to foster interest and engagement in learning within students outside medical education.[8,9] It is not surprising, therefore, that deliberate reflection led to longer engagement in the learning activity. This positive effect of deliberate reflection on actual engagement in

studying the learning material is also consistent with previous findings showing that deliberate reflection fostered medical students' interest in knowing more about the cases.[7]

As deliberate reflection increased both study time and test scores, the latter may have been a consequence of the former. By leading to investment of more time in studying the learning material, deliberate reflection increased learning outcomes. This assumption is consistent with the positive relationship between the amount of study time and learning results observed in experiments in which students worked individually,[13,14] as our participants did, such as in research on a teaching journal's quality assessment. Engagement in learning, after all, has been shown to be a consistent and important mediator of learning, even if the time engagement is short, a matter of minutes,[12] such as what is expected to happen when physicians or medical students are confronted with an uncertain diagnosis. However, another possible explanation for our findings has to be considered: deliberate reflection may have mobilised students' prior knowledge and, once activated, it would make it easier to understand the new information presented in the study material and to integrate it into existing knowledge structures.[22,23] It could also be an interaction between these two effects, but exploring the mechanisms underlying the effects of reflection was not within the scope of our study and requires further research.

Educational implications

Our findings add empirical support to the importance of deliberate reflection as a key competence for medical students and practitioners. It demonstrates the potential of deliberate reflection upon clinical cases, a simple, short-time-consuming process of confronting diagnostic

hypotheses with patients' clinical data, as a motivational force for students' engagement in learning activities. Clinical teachers could use it as a strategy to motivate their students to study medical topics related to clinical cases. It also expands the evidence of engagement in learning as an important mediator of learning outcomes in the medical domain. To the best of our knowledge, our findings provide the first empirical evidence of the effects of deliberate reflection on study engagement in learning and learning outcomes in medical education.

Limitations

Our study was run in a single medical school, with participants in the same year of training and who worked with a single clinical topic, jaundice, which limits the generalisability of its findings. We measured students' engagement by time on task because it is a variable known to influence learning outcomes,[12] but there could be other dimensions, such as intellectual engagement, that we did not explore. Also, because we used participants' immediate post-test scores as a measure of learning outcome, we cannot foresee if the positive effect of deliberate reflection that we observed on learning would last longer. Finally, the task of solving crosswords in the control condition may be seen as potentially harmful for the learners, as the crosswords could have distracted them from the diagnostic task. However, because students in the control group provided alternative diagnoses for the case *before* rather than *after* solving the crossword, they would not be affected by the crossword task *while* generating the alternative diagnoses. The crossword may eventually have influenced the accuracy of their final diagnosis, but this was not an outcome measurement of the study. Our intention was to compare the effect of deliberate reflection and providing alternative diagnoses on engagement in a subsequent study task. Nevertheless, it cannot be excluded that solving the crossword generated excessive cognitive load, which somehow carried over to the study task, although it is worth noticing that the deliberate reflection procedure has also been shown to involve high cognitive load.[24]

CONCLUSIONS

In summary, we studied the influence of deliberate reflection while working with to-be-diagnosed clinical cases on engagement in learning and

learning outcomes among fourth-year medical students. We found positive effects of deliberate reflection on both, which adds evidence to sustain it as an important competence for medical students. It also expands the evidence on the potential benefits of a deliberate reflection procedure that can be easily used by clinical teachers to motivate their students, possibly in both real and simulated clinical environments. Nevertheless, there are still relevant questions. How did deliberate reflection foster learning? Through expansion or reorganisation of knowledge? Would the observed effect of deliberate refection be similar in different contexts, for example, with more or less experienced students? The answers can contribute to better practices in medical schools. Paraphrasing Dewey, it is necessary to climb the field's tree, survey additional facts and see how these things relate to one another.[1]

REFERENCES

1 Dewey J. What is thought? In: Dewey J, ed. *How We Think*. Boston, MA: D. C. Heath and Co 1910;2–22.
2 Schön DA. Understanding the need for artistry in professional education. In: Schön DA, ed. *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in Professions*. San Francisco, CA: Jossey-Bass Publishers 1987;1–40.
3 Sandars J. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach* 2009;**31** (8):685–95.
4 Ericsson KA. Acquisition and maintenance of medical expertise: a perspective from the expert-performance

approach with deliberate practice. *Acad Med* 2015;**90** (11):1471–86.

5 Mamede S, Schmidt HG. Correlates of reflective practice in medicine. *Adv Health Sci Educ* 2005;**10** (4):327–37.

6 Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ* 2008;**42** (5):468–75.

7 Ribeiro LMC, Mamede S, Moura AS, de Brito EM, de Faria RMD, Schmidt HG. Effect of reflection on medical students' situational interest: an experimental study. *Med Educ* 2018;**52** (5):488–96.

8 Rotgans JI, Schmidt HG. Situational interest and learning: thirst for knowledge. *Learn Instruction* 2014;**32**:37–50.

9 Rotgans JI, Schmidt HG. Situational interest and academic achievement in the active-learning classroom. *Learn Instruction* 2011;**21** (1):58–67.

10 Linnenbrink-Garcia LP, Patall EA, Messersmith EE. Antecedents and consequences of situational interest. *Br J Educ Psychol* 2013;**83**:591–614.

11 Hidi S. Interest and its contribution as a mental resource for learning. *Rev Educ Res* 1990;**60** (4):549–71.

12 Fredrick WC, Walberg HJ. Learning as a function of time. *J Educ Res* 1980;**73** (4):183–94.

13 Glogger-Frey I, Fleischer C, Gruny L, Kappich J, Renkl A. Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learn Instruction* 2015;**39**:72–87.

14 Ainley M, Hidi S, Berndorff D. Interest, learning and the psychological processes that mediate their relationship. *J Educ Psych* 2002;**94** (3):545–61.

15 Gettinger M. Time allocated and time spent relative to time needed for learning as determinants of achievement. *J Educ Psych* 1985;**77** (1):3–11.

16 Mamede S, van Gog T, Moura AS, de Faria RMD, Peixoto JM, Rikers RM, Schmidt HG. Reflection as a strategy to foster medical student's acquisition of diagnostic competence. *Med Educ* 2012;**46** (5):464–72.

17 Mamede S, van Gog T, Moura AS, de Faria RMD, Peixoto JM, Schmidt HG. How can students' diagnostic competence benefit most from practice with clinical cases? The effects of structured reflection on future diagnosis of the same and novel diseases. *Acad Med* 2014;**89** (1):121–7.

18 Meyer BJF. The organization of prose and its effects on memory. In: van Dijk TA, Hendricks WO, eds. *North-Holland Studies in Theoretical Poetics*, vol 1. Amsterdam/Oxford: North-Holland Publishing Co. New York, NY: American Elsevier Publishing Co. 1975;249.

19 Schiefele U, Krapp A. Topic interest and free recall of expository text. *Learn Indiv Diff* 1996;**8** (2):141–60.

20 Fraenkel JR, Wallen NE, Hyun H. Content analysis. In: Fraenkel JR, Wallen NE, Hyun H, eds. *How to Design and Evaluate Research in Education*, 8th edn. New York: McGraw-Hill 2012;477–504.

21 Sullivan GM, Feinn R. Using effect size - or why the p value is not enough. *J Grad Med Educ* 2012;**4** (3):279–82.

22 Mayer RE. Information processing. In: Harris KR, Graham S, Urdan T, eds. *APA Educational Psychology Handbook: Vol. 1. Theories, Constructs, and Critical Issues*, 1st edn. Washington, DC: American Psychology Association 2012;85–99.

23 Schmidt HG, Devolder ML, Degrave WS, Moust JHC, Patel VL. Explanatory models in the processing of science text - the role of prior knowledge activation through small-group discussion. *J Educ Psych* 1989;**81** (4):610–9.

24 Ibiapina C, Mamede S, Moura A, Elói-Santos S, van Gog T. Effects of free, cued and modelled reflection on medical students' diagnostic competence. *Med Educ* 2014;**48**:796–805.