Exploring mechanisms underlying learning from deliberate reflection

Published in:
Medical Education

Publication status and date:
Published: 01/03/2021

DOI (link to publisher):
10.1111/medu.14410

Document Version
Publisher's PDF, also known as Version of record

Citation for the published version (APA):
Ribeiro, L. M. C., Mamede, S., de Brito, E. M., Moura, A. S., de Faria, R. M. D., & Schmidt, H. G. (2021). Exploring mechanisms underlying learning from deliberate reflection: An experimental study. *Medical Education*, *55*(3), 404-412. https://doi.org/10.1111/medu.14410

Received: 12 August 2020 | Revised: 13 October 2020 | Accepted: 3 November 2020

DOI: 10.1111/medu.14410

CLINICAL REASONING

Exploring mechanisms underlying learning from deliberate reflection: An experimental study

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

[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, Rotterdam, The Netherlands

[3]Department of Popaedeutics, Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil

[4]Education and Research Center, Santa Casa BH, Belo Horizonte, Brazil

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

Funding information
José do Rosário Vellano University (UNIFENAS) supported the study by providing Uslaene Rocha de Lima, research assistant, with a one-year scientific initiation scholarship.

Abstract

Introduction: Previous research suggests that, relative to generating a differential diagnosis, deliberate reflection during practice with clinical cases fosters learning from a subsequently studied scientific text and promotes interest in the subject matter. The present experiment aimed to replicate these findings and to examine whether motivational or cognitive mechanisms, or both, underlie the positive effects of reflection.

Methods: A total of 101 5th-year medical students participated in an experiment containing four phases: Students (a) diagnosed two clinical cases of jaundice-related diseases either through deliberate reflection or differential diagnosis; (b) reported their situational interest and awareness of knowledge gaps; (c) studied a text about jaundice, either under free or restricted time; and (d) recalled the text. Outcome measures were text-recall, situational interest and awareness of knowledge gaps.

Results: A main effect of diagnostic approach on recall of the text was found, with the reflection group recalling more studied material than the differential diagnosis group (means: 72.56 vs 58.80; $P = .01$). No interaction between diagnostic approach and study time (free or restricted) emerged, nor was there a main effect of the latter. Relative to the differential diagnosis group, students who reflected upon the cases scored significantly higher on both situational interest (means: 4.45 vs 3.99, $P < .001$) and awareness of knowledge gaps (means: 4.13 vs 3.85, $P < .01$).

Discussion: Relative to generating differential diagnoses, reflection upon clinical cases increased learning outcomes on a subsequent study task, an effect that was independent of study time, suggesting that cognitive mechanisms underlie this effect, rather than increases in motivation to study. However, higher scores on situational interest and awareness of knowledge gaps and a tendency towards larger gains when time was free suggest that higher motivation may also contribute to learning from reflection.

1 | INTRODUCTION

Reflection can be defined as the processes of questioning one's own assumptions and their consequences for decision making. It is usually triggered and is particularly useful when we face challenging situations that raise uncertainty.[1] Medical education places much value in developing medical students' ability to engage in reflection because as professionals they will manage ill-defined problems, such as clinical cases for which a clear diagnosis is not available. Indeed, a series of studies have shown that engaging in deliberate reflective processes can improve the accuracy of diagnostic reasoning, as detailed below. Besides being helpful in professional practice, however, reflection is also considered important as a learning tool. It is expected to foster students' engagement in learning activities,[2,3] and there is some empirical evidence that reflecting upon to-be-diagnosed clinical cases can foster medical students' learning from subsequent study activities relevant to those cases.[4,5] However, this evidence is still preliminary and the mechanisms that underlie such potential positive effects of reflection are not yet understood. The experiment reported in this article examined whether the previously observed positive effect of reflection on medical students' learning replicates and explored two possible mechanisms that may underlie this effect: motivating students to engage in learning activities and fostering cognitive processes that facilitate learning.

The aforementioned studies on the effects of reflection on students' learning[4,5] employed the 'deliberate reflection procedure', originally developed as a tool to improve physicians' performance.[6] The procedure involves reasoning through clinical problems by generating a tentative diagnostic hypothesis, confronting this hypothesis with the patient's clinical findings, considering alternative explanations for the problem, and arguing for and against each emerging explanation before making a final decision. Besides helping physicians solve difficult clinical cases,[6] the deliberate reflection procedure has been used as a learning tool. Mamede et al[7,8] observed that students who engaged in deliberate reflection during practice with clinical cases provided better diagnoses of new cases of the same (or related) diseases in the future than their peers who had used more conventional approaches. The results were attributed to a refinement of students' mental representation of the diseases, since no additional knowledge was offered to the students in these experiments.

Ribeiro et al[4,5] used the same deliberate reflection procedure in two experiments with medical students, exploring the effects of reflection on learning of *new* material (as opposed to refinement of previously learned knowledge). Specifically, they studied the potential of reflection to foster students' engagement with, and learning from, a text that provided detailed explanations about the cases. In the first experiment, fourth-year medical students diagnosed clinical cases either through the deliberate reflection procedure or through generating differential diagnoses. Scores on students' situational interest (SI), as measured by a self-reported questionnaire,[9] were obtained. SI is a form of transitory interest also labelled 'thirst' for knowledge.[10] SI is expected to possibly derive from the awareness of knowledge gaps (AKG) that emerges when we try to solve challenging problems; the feeling of being unable to make sense of things can trigger a need to close those gaps. Indeed, SI has been shown to be a good predictor of engagement in learning activities.[9-12] The experiment showed deliberate reflection to foster students' SI relative to giving differential diagnosis, but did not evaluate whether students actually engaged in more learning activities related to the cases. A follow-up experiment showed that, relative to differential diagnosis, deliberate reflection fostered engagement with a subsequently studied relevant text (as measured by the amount of time spent processing that text) and higher cued recall from this text.[5] These findings suggest that deliberate reflection is an effective learning tool able to foster engagement in learning activities and, ultimately, learning. The experiments, however, did not investigate the psychological mechanisms underlying the positive effect of reflection on students' learning. Since time spent in learning activities can influence learning outcomes,[13] the observed benefits to learning could have arisen from greater motivation leading to longer engagement with the text. That means, deliberate reflection would foster learning simply by increasing time invested in studying the material. Reflection could, however, also have affected the nature of the information processing thereby directly influencing students' cognition (ie the knowledge they took away from studying the text). It would be reasonable to expect that reflection would foster cognitive processes such as activation of prior knowledge and/or elaboration on the new information, as it has been observed in other learning strategies involving problem-solving.[14-19]

We conducted an experiment to examine whether increased *motivation*, facilitation of *cognitive processing*, or a combination of both mechanisms underlie the effect of deliberate reflection on learning. Fifth-year medical students diagnosed two clinical cases either through deliberate reflection or by giving differential diagnosis. Subsequently, they performed a study task, reading a relevant text, either under free- or restricted-time conditions, and took a cued recall test. Assume for a moment that *only* higher levels of motivation (as expressed in longer engagement with the text) would cause an effect of diagnostic approach on cued recall. Then, a main effect of diagnostic approach and an interaction effect between diagnostic approach and time (free vs restricted) would emerge because under restricted-time conditions no difference in cued recall would be observed. In addition, a main effect of time (free vs restricted) would be expected. However, if the positive effect of deliberate reflection was the result of better cognitive processing *only*, a main effect of diagnostic approach on cued recall would be expected, and, more importantly, *no* interaction with study time (free vs restricted) should be seen. The latter would signify that the effect of the treatment is the same, even if study time is restricted (ie even when motivation is obstructed from expressing itself through longer engagement). Third, presence of a main effect of both time and diagnostic approach, and absence of an interaction effect, would suggest that both processes, motivational and cognitive, have been involved. To further explore the possible motivational mechanism underlying reflection, participants' SI and AKG were also measured.

We hypothesised that, relative to students who gave differential diagnoses, students who reflected upon the cases would have higher scores on SI and AKG.

2 | METHODS

2.1 | Design

The experiment consisted of an exercise with four phases that took place sequentially in a single session: (1) a diagnostic phase, (2) measurement of SI and AKG, (3) study of a relevant text and (4) a cued recall test. In the diagnostic phase, participants diagnosed two clinical cases with jaundice as the chief complaint either by following the deliberate reflection procedure or giving differential diagnosis. After diagnosing each case, participants answered questionnaires on SI and AKG. In the subsequent study phase, participants studied the same text about the differential diagnosis of jaundice either with free time or within a fixed maximum amount of time allocated for the task. Finally, all participants performed a cued recall test about the material they had just studied. There were, therefore, four different experimental conditions, with the approach used for the diagnostic phase (deliberate reflection vs differential diagnosis) crossed with the time conditions for the study phase (free study time vs restricted study time). Participants were randomly allocated to one of the four conditions (see Figure 1).

2.2 | Setting and participants

Three-hundred and fifty-fifth-year medical students at the Federal University of Minas Gerais (UFMG), in Belo Horizonte, Brazil, were invited to volunteer for the experiment between July 2018 and July 2019. This school has a six-year curriculum with the two final years dedicated to clerkships. We recruited these students because at this time of their training they have been exposed to knowledge about jaundice but have limited clinical experience with it. The experiment was an extracurricular activity carried out in different sessions to accommodate students' timetables. Participants gave their written consent and were offered the possibility to discuss the experiment with the authors after completing it.

2.3 | Materials

All materials were presented to participants via computer using the *Qualtrics* suite. To begin, participants were asked to rate their confidence in diagnosing patients with jaundice using a five-point scale (very low confidence to very high confidence) and to estimate how many real patients with jaundice they recalled having encountered (either 0, 1-3, >3).

Phase 1 used two clinical cases that were employed in previous studies.[4,5] Both had jaundice as the main clinical finding and consisted of written descriptions of clinical symptoms, physical examination and laboratory tests. Each had a confirmed most likely diagnosis: acute viral hepatitis or choledocholithiasis.[7,8] We selected jaundice because it allowed for an exercise that covered relevant clinical information in a condensed time, which usually fosters participation and compliance with instructions.

Phase 2 used questionnaires on SI and AKG that had similarly been used in previous research.[4] Rotgans and Schmidt[9,10,20,21] developed the SI questionnaire, which was shown to have acceptable reliability when administered to medical students at a similar level of training.[4] The SI questionnaire contains six questions like *I was totally focused while working on this task*. The AKG questionnaire was developed by Glogger-Frey et al[22] for research on instructional design and has similarly shown good reliability in previous study with medical students.[4] The questionnaire has nine questions such as *working on this task revealed I don´t know certain things yet*. The questionnaires requested participants to answer each question by using five-point Likert scales, ranging from *not true at all* to *very true to me*, and are available as Supplementary Material S1.

In Phase 3, the study material consisted of an illustrated text presenting a review of bilirubin physiopathology followed by the presentation of the clinical cases participants had just diagnosed, with the key clinical findings valuable to differentiate between the causes

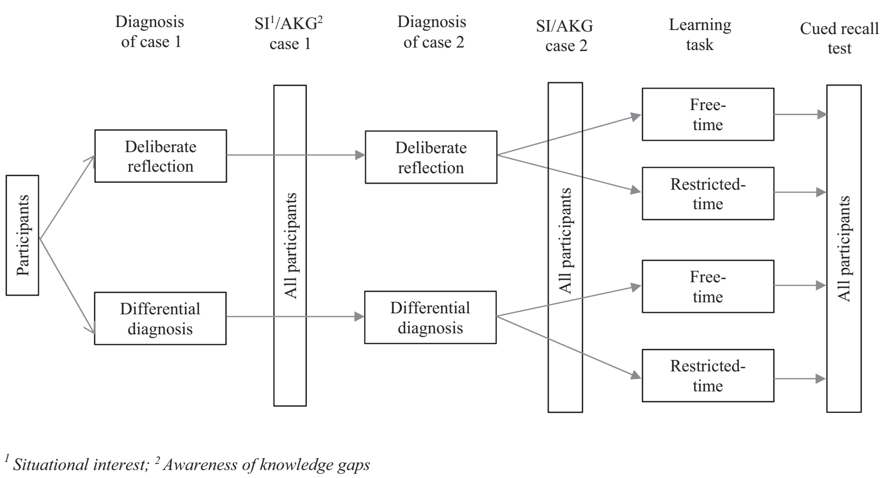

FIGURE 1 Experiment design and tasks flow [Colour figure can be viewed at wileyonlinelibrary.com]

of jaundice highlighted and linked to boxes with their interpretation and explanation (eg Coluria indicates cholestasis because free direct bilirubin is excreted in urine).

Finally, Phase 4 used a cued recall test with eight open-ended questions to measure students' learning of the study material. The test, previously used in an experiment with similar participants,[5] addressed topics on clinical history, physical examination and laboratory tests concerning the differential diagnosis of jaundice. Each item cued recall of a specific part of the material and requested the participants to write down information from the text they had just read. For example, 'List all the relevant physical examination findings to the evaluation of patients with jaundice and explain how they help on the differential diagnosis'. The cued recall test is available as Supplementary Material S2.

2.4 | Procedures

The experimental interventions were presented during Phase 1 (manipulation of the diagnostic instructions) and Phase 3 (time allowed for study).

For Phase 1, students in the deliberate reflection condition were requested to read the case and follow a procedure to reflect upon the case[6] which, briefly requires (a) typing the most likely diagnosis for the case, (b) typing the clinical findings that are in line with their first diagnostic hypothesis; findings that contradict it; and those that are expected (if this first hypothesis were true) that are not described in the case; (c) typing two alternative diagnoses and the same findings for each diagnosis; and (d) drawing a conclusion on the most likely diagnosis for the case. Students in the differential diagnosis condition were asked to carry out the following sequence of steps: (a) read the case, type the most likely diagnosis and two alternative diagnoses, (b) work on a crossword containing medical terms not related to the cases (to ensure similar time spent in the two conditions while minimising this group's engagement in reflection) and (c) typing their conclusion on the most likely diagnosis for the case. The instructions stated that, although it could seem irrelevant to complete a crossword, a task like it might help decision making and should, therefore, be taken seriously. For all groups, time to progress throughout the sequence of steps in the diagnostic phase was controlled by Qualtrics as follows: 2 minutes to type the most likely diagnosis/diagnoses; 7 minutes to reflect upon the case (reflection condition) or to solve the crossword and type the final diagnosis (control condition).

After diagnosing the first clinical case, the computer moved participants to Phase 2 (the SI and AKG questionnaires). Participants answered the questionnaires at their own pace. Next, the suite presented the second clinical case followed by a second set of SI and AKG questionnaires.

Subsequently, in Phase 3 (the learning task), the study material was presented with a request that students read it. Time was controlled by the software with participants being told whether study time would be restricted or not. Those for whom study time was restricted had up to 180 seconds to read the text. This was the median time a similar group of students spent on the same material in a previous experiment in which no restriction of time was imposed.[5] Participants could move onto Phase 4 in <180 seconds if they wanted to. Those for whom time to study was free could read it as long as they wanted.

Upon completion of the learning task, the next screen presented participants with Phase 4, the cued recall test. They were requested to answer the questions by making an effort to write all information they could remember. For the test phase, there was no restriction of time.

The experiment was conducted in a computer laboratory with students working individually. Each student received a copy of a link that would give access to one of the four versions of the Qualtrics experiment. The links had been previously organised in random order to randomly assign participants to one of the four conditions, since it was not possible to anticipate which students would volunteer for the experiment.

Participants could neither skip steps nor move backward on the exercise. The software automatically controlled and recorded time spent on each phase and recorded each participant's responses. No information about their diagnostic performance was provided to participants while they worked on the cases, but the study material of Phase 3 explained the correct diagnoses. One of the authors was present in all sessions to help in case of computer problems and to inhibit students from consulting resources other than those presented as part of the exercise.

2.5 | Data analysis

The diagnoses provided by the students in the diagnostic task were independently assessed by two board-certified internists (LMCR and EMB), who were blinded to experimental condition. They classified each response as correct (scored 1) whenever the core diagnosis was present (eg '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 (eg 'biliary colic' in the 'choledocholithiasis' case); and incorrect (scored 0) when the response did not fall into any of these categories. The raters agreed in the score attributed in 87% of the responses and resolved discrepancies by 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 by summing, for each participant, scores achieved on the two cases and subsequently averaging within each experimental condition.

The amount of information reported in participants' responses (cued recall task) was evaluated by counting the number of idea units present in each response.[23,24] We considered idea units to be text fragments, such as a word or a short sentence, meaningful to the task. Each idea unit was scored with one point as long as it met three criteria: consisted of correct information, was consistent with the question cue and was actually present in the study material. For example, for the cue 'Explain how prothrombin activity helps differentiating the causes of jaundice' one idea unit was 'prothrombin activity depends on liver function'. Two authors (LMCR and EMB), independently assessed 10% of participants' responses, reaching an initial agreement of 84%. After discrepancies were resolved, the first author coded all participant responses. Only manifest content was considered (ie no possible underlying meanings were counted).[25] Supplementary Material S2 presents an example of the test scoring procedure.

The main outcome measurements of the study were: Cued recall (as measured by the total number of accurate idea units produced in cued recall), and SI and AKG, taken as measures of engagement in learning. To compute cued recall, we first summed the number of idea units recorded by each participant and then averaged across participants for each experimental condition. To obtain the SI and AKG scores, for each participant, we averaged the ratings provided to the items of the questionnaires, and subsequently computed the mean for each experimental condition. Actual processing times were registered by the software for each participant and we computed means for each experimental condition.

Data were analysed using SPSS for Macintosh, version 25 (IBM Corp). Significance level was set at $P < .05$. To check whether groups were similar in terms of the extraneous variables, age, initial diagnostic accuracy, and confidence in diagnosing jaundice, two-way ANOVAs were conducted with diagnostic approach and study time (free vs restricted). Pearson *chi*-square tests were carried out comparing gender and previous experience with patients with jaundice between the four experimental groups. All other variables were analysed applying ANOVAs: separate one-way tests to compare SI and AKG according to diagnostic approach and a two-way test with diagnostic approach and study time (free vs restricted) on cued recall.

3 | RESULTS

3.1 | Inclusion of participants and outliers

One-hundred and five (30%) of the 350 invited students agreed to participate. Four outliers (*z*-scores > 2.58 in study time and/or test-score) were removed after exploratory data analysis; this is a commonly applied threshold as only 1% of values can be expected to be outside this range.[26] This led to 101 participants being included in the analyses.

3.2 | Analysis of success of randomisation

Table 1 presents participants' characteristics as a function of experimental condition. No significant differences between groups emerged in age, $F(1,97) = 0.88$, $P = .17$, gender, $X^2(3) = 2.74$, $P = .43$, previous experience with patients with jaundice, $X^2(3) = 0.93$, $P = .82$, confidence in diagnosing jaundice $F(1,97) = 0.81$, $P = .37$, or initial diagnostic accuracy, $F(1,97) = 0.18$, $P = .67$.

3.3 | Analysis of outcomes

Table 2 presents the main outcome measurements. There was a significant main effect of diagnostic approach on cued recall [means: 72.56 vs 58.80, $F(1, 97) = 6.43$, $P = .01$, $d = 0.54$], with students who deliberately reflected upon the cases showing higher scores on the test than those who gave differential diagnoses. Importantly, there was no effect of study time (free vs restricted) on cued recall [$F(1, 97) = 0.24$, $P = .63$, *ns*]; nor was an interaction effect demonstrated between diagnostic approach and study time [$F(1, 97) = 0.65$, $P = .42$, *ns*]. These findings support a cognitive explanation for the effect of

TABLE 1 Characteristics of the participants as a function of diagnostic approach and study time (standard deviation in brackets)

	Free study time		Restricted study time		
	Deliberate reflection	Differential diagnosis	Deliberate reflection	Differential diagnosis	Total
N	25	24	27	25	101
Age	23.96 (2.35)	23.5 (2.04)	23.33 (1.8)	24 (2.3)	23.7 (2.12)
Gender					
Male	15	14	15	19	63
Female	10	10	12	6	38
Previous experience					
No experience	8	4	5	4	21
1-3 patients	14	17	20	17	68
>3 patients	3	3	2	4	12
Confidence in diagnosis (range 1-5)	2.88 (0.73)	2.79 (0.78)	2.89 (0.80)	3.08 (0.81)	2.91 (0.78)
Initial diagnostic accuracy (range: 0-2)	1.32 (0.78)	1.41 (0.71)	1.61 (0.56)	1.60 (0.47)	1.49 (0.64)

deliberate reflection on learning and cued recall, at the expense of a motivational, or a combined cognitive/motivational explanation (see also Figure 2).

As in a previous study,[4] a significant main effect of diagnostic approach on SI was found, with participants who diagnosed the cases through deliberate reflection showing higher scores on SI [means: 4.45 vs 3.99, $F(1,99) = 26.27$, $P < .001$, $d = 1.01$] than those who gave differential diagnoses. A main effect in favour of reflection also emerged for AKG [means: 4.13 vs 3.85, $F(1,99) = 8.42$, $P < .01$, $d = 0.58$]. These findings suggest that deliberate reflection fosters an interest in the topic-at-hand and makes students aware of knowledge deficits. This enhanced interest does not seem to translate into increased engagement with the subject matter as measured by study time. We will return to this issue in the Discussion section.

As a post hoc test to further explore the influence of time, we used structural equation modelling to assess whether actual study time served as a moderator variable between diagnostic approach and cued recall for participants in the free study time condition. Figure 3 presents a model of this relationship that fits the data, but the beta-weight of the relationship between diagnostic approach and study time is not significantly different from zero.

4 | DISCUSSION

In this experiment, we investigated whether deliberate reflection while practising with clinical cases fosters medical students' learning and examined two potential mechanisms through which deliberate reflection could affect learning: by fostering motivation and by facilitating cognitive processes involved in learning. To do so, we ran an experiment in which fifth-year medical students diagnosed clinical cases presenting patients with jaundice either through deliberate reflection or differential diagnosis, reported their SI and AKG about the topic, studied a text referring to the topic either with free or restricted time and, finally, answered a test about this text. Students who deliberately reflected upon the cases reported higher scores on SI and AKG and obtained higher scores on the cued recall test relative to those who gave differential diagnosis. These findings are in line with our hypotheses that working on cases through deliberate reflection would be more beneficial to students' learning than the more conventional approach of generating a differential diagnosis, replicating what was observed in previous experiments.[4,5] The effect size of deliberate reflection was large for SI and medium for AKG and cued recall.[27]

This effect of diagnostic approach on cued recall was independent of study time (free vs restricted), given that there was neither a significant effect of study time on cued recall nor an interaction between diagnostic approach and study time. These findings suggest a cognitive processing only explanation for the findings, rejecting the possibility of motivational influences. This argument is outlined in the Introduction: if a motivational mechanism-only would underlie the effect, keeping time restricted would wipe-out any differences between both diagnostic approaches under that condition, causing an interaction effect in the data. We did not observe such interaction, making an explanation in terms of motivation-only unlikely. A third possibility that we considered was that *both* cognition and motivation play a role in the effect of deliberate reflection on learning. However, this possibility was falsified by the fact that there was no main effect of time (free vs restricted) on cued recall. Were a combination of both operating, performance on the cued recall test should be significantly higher under free-time conditions than under restricted time. We will return to the issue of motivation below.

Thus, deliberate reflection upon the cases seems to have facilitated cognitive processes that are known to influence learning outcomes. One such cognitive process is the *activation of prior knowledge*, as it has been observed in the context of PBL.[14-17] Although we did not have measures of information processing while students diagnosed the cases, it is reasonable to think that confronting diagnostic hypotheses with patients' data in a systematic way mobilised

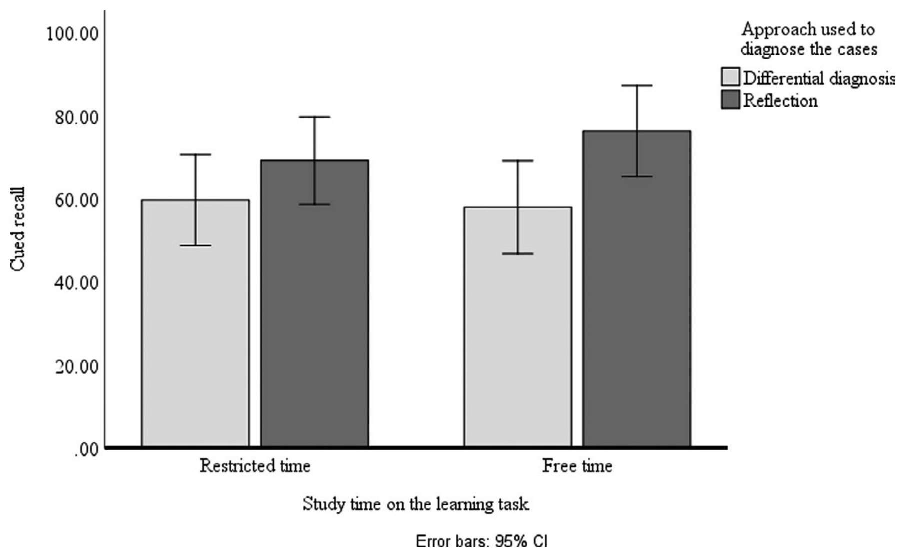

FIGURE 2 Mean cued recall as a function of diagnostic approach and study time (free vs restricted) [Colour figure can be viewed at wileyonlinelibrary.com]

students' knowledge from long-term to working-memory[28] more extensively than having them listing differential diagnoses. Besides the activation of prior knowledge, by requiring learners to connect clinical findings with each other and with the hypotheses, reflection might also have facilitated *elaboration* of knowledge. These processes fostered by reflection possibly facilitated processing and incorporation of the new information encountered in the learning task. This result is consistent with studies on PBL and self-explanation that have found students who attempt to explain, in their own words, the problems with which they are working to perform better on tests on knowledge relevant to those problems than those who do not explain.[15,29] In summary, deliberate reflection is suggested to activate prior knowledge and elaboration to a larger extent than simply listing differential diagnoses, thereby facilitating processing of new information and subsequent recall to a larger extent.

A tantalising finding, however, is the unequivocal positive effect of deliberate reflection on two measures of engagement, situational interest (SI) and awareness of knowledge gaps (AKG), often considered measures of motivation.[9-12,20,21] Deliberate reflection compels students to scrutinise different diagnostic hypotheses for the cases, arguing for and against them before making any conclusions, which may create additional challenge and bring knowledge gaps to surface, thereby stimulating interest in the topic-at-hand.[10] We expected increased interest to translate into longer engagement with the learning task for those whose study time was not restricted. Students in the deliberate reflection condition indeed tended to invest more time than their colleagues from the differential diagnosis condition, but the expected positive interaction between diagnostic approach and study time did not emerge. Text-reading research, however, shows that interest is not associated only with study time, but also with choices, and can translate into what people choose to focus attention on.[11,24] This is a motivational mechanism that might have influenced results particularly in the restricted-time groups, who were aware their time would be restricted and might, therefore, have focused attention on information that could fulfil their perceived knowledge gaps, optimising study time. As our experiment does not allow us to identify specific motivational mechanisms, these assumptions require further investigation.

A second finding that does not seem to entirely fit with our conclusion that only cognitive processes can be held responsible for the effects of deliberate reflection and that study time, as an indicator of engagement, does not play a role, is that, in the free-time condition, students who reflected spent 17% more time on reading the text. Although this difference was not significant, we decided to relate, for the free-time condition only, actual study time as a moderator variable between diagnostic approach and cued recall, in a post hoc analysis using structural equations modelling. Figure 3 presents a model of this relationship that fits the data. However, the beta-weight of the relationship between diagnostic approach and study time is non-significantly different from zero. This suggests that study time plays a role in determining performance on the cued recall test, but this effect is *not* driven by the diagnostic approach. Rather, it must be driven by extraneous factors, possibly related to individual differences among students.[30]

4.1 | Limitations

This experiment was run in a single medical school, with all participants in the same training year, and a single medical subject, jaundice, was explored. Though reasons that would make the findings specific to these particular students or topic are not clear, the generalisability of the findings may have been compromised as a result. In addition, we cannot exclude the possibility that participants in the differential diagnosis condition who had to solve crosswords puzzles may have experienced cognitive load that somehow carried over to the test phase and could have negatively influenced their performance. However, deliberate reflection has been shown to involve substantially higher cognitive load than other problem-solving approaches with which it has been compared.[31] Therefore, it is unlikely that cognitive load has affected only participants in the differential diagnosis condition. Finally, since we used immediate post-test scores as a measure of learning outcome, we are unable to know if the positive effect of deliberate reflection on learning would last longer.

TABLE 2 Mean cued recall (in number of idea units), actual time spent on learning (in seconds), SI, and AKG as a function of diagnostic approach and study time (standard deviation into brackets)

	Free study time		Restricted study time		
	Deliberate reflection	Differential diagnosis	Deliberate reflection	Differential diagnosis	Total
N	25	24	27	25	101
Cued recall	76.24 (32.39)	57.91 (28.94)	69.14 (22.21)	59.64 (26.07)	65.88 (28.11)
Actual time[a]	405.09 (222.89)	347.17 (183.60)	168.80 (23.39)	170.02 (20.04)	269.97 (176.55)
SI	4.47 (0.44)	4.01 (0.49)	4.44 (0.44)	3.97 (0.52)	4.23 (0.51)
AKG	4.24 (0.41)	3.86 (0.55)	4.03 (0.41)	3.84 (0.42)	3.99 (0.49)

[a]Actual time needed to process the text, although represented here, was not further analysed because half of the groups worked under time restrictions, rendering the resulting data of limited value. However, actual study time acquired under the free study time condition was further post hoc analysed.

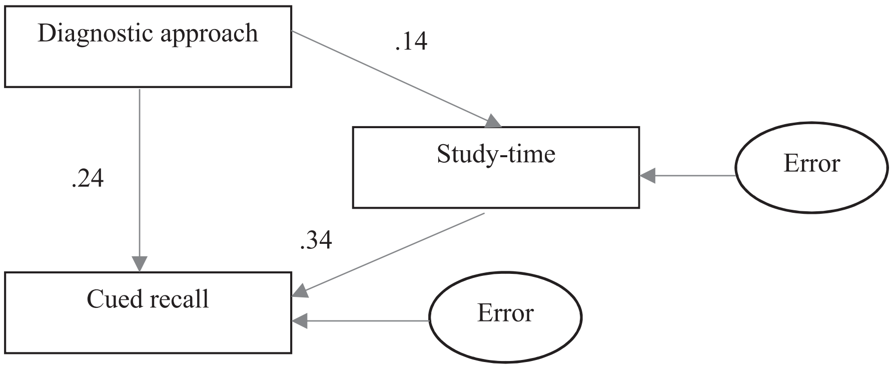

FIGURE 3 Structural equations model of the relationship between diagnostic approach, study time, and cued recall for the free-time condition only (N = 49) [Colour figure can be viewed at wileyonlinelibrary.com]

5 | CONCLUSION

Our experiment provides additional evidence of the positive effect of deliberate reflection on medical students' learning of subsequent relevant material observed in a previous experiment[5] and suggests that this effect is mediated by cognitive mechanisms. Future research should study how these mechanisms are facilitated by reflection, thereby opening the door for an optimal use of deliberate reflection by clinical teachers.

ACKNOWLEDGEMENTS

The authors are grateful to Uslaene Rocha de Lima, research assistant, and to the students who dedicated their time and efforts to this study. They are also thankful to Dr Inga Glogger-Frey and Prof. Dr Alexander Renkl who kindly shared their questionnaire on awareness of knowledge gaps and allowed the authors to use it in the experiment.

AUTHORS' CONTRIBUTIONS

Ligia Maria Cayres Ribeiro had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Ligia Ribeiro, Silvia Mamede, Henk Schmidt. Acquisition of data: Ligia Ribeiro, Alexandre Moura, Eliza Brito, Rosa Faria. Statistical analysis: Ligia Ribeiro, Silvia Mamede. Analysis and interpretation of data: Ligia Ribeiro, Silvia Mamede, Eliza Brito, Alexandre Moura, Rosa Faria, Henk Schmidt. Drafting of the manuscript: Ligia Ribeiro. Critical revision for the manuscript for important intellectual content: Ligia Ribeiro, Silvia Mamede, Eliza Brito, Alexandre Moura, Rosa Faria, Henk Schmidt. Administrative, technical or material support: Silvia Mamede, Henk Schmidt, Rosa Faria. Study supervision: Silvia Mamede, Henk Schmidt.

ETHICAL APPROVAL

Ethical approval for the study was provided by Research Ethics Committee of Federal University of Minas Gerais (on April 17, 2018, number 2.604.504).

ORCID

Ligia Maria Cayres Ribeiro https://orcid.org/0000-0002-4277-3066
Silvia Mamede https://orcid.org/0000-0003-1187-2392
Henk G. Schmidt https://orcid.org/0000-0001-8706-0978

REFERENCES

1. Dewey J. *How we Think*. Boston: D. C. Haeth and Co.; 1913.
2. Sandars J. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach*. 2009;31(8):685-695.
3. Uygur J, Stuart E, De Paor M, et al. A Best Evidence in Medical Education systematic review to determine the most effective teaching methods that develop reflection in medical students: BEME Guide No. 51. *Med Teach*. 2019;41(1):3-16.
4. 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):488-496.
5. Ribeiro LMC, Mamede S, de Brito EM, Moura AS, de Faria RMD, Schmidt HG. Effects of deliberate reflection on students' engagement in learning and learning outcomes. *Med Educ*. 2019;53(4):390-397.
6. Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnosis. *Med Educ*. 2008;42(5):468-475.
7. Mamede S, van Gog T, Moura AS, et al. Reflection as a strategy to foster medical student's acquisition of diagnostic competence. *Med Educ*. 2012;46(5):464-472.
8. 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):121-127.
9. Rotgans JI, Schmidt HG. Situational interest and academic achievement in the active-learning classroom. *Learn Instr*. 2011;21(1):58-67.
10. Rotgans JI, Schmidt HG. Situational interest and learning: Thirst for knowledge. *Learn Instr*. 2014;32:37-50.
11. Ainley M, Hide S, Berndorff D. Interest, learning and the psychological processes that mediate their relationship. *J Educ Psychol*. 2002;94(3):545-561.
12. Linnenbrink-Garcia L, Pattal EA, Messersmith EE. Antecedents and consequences of situational interest. *Br J Educ Psychol*. 2013;83(4):591-614.
13. Fredrick WC, Walberg HJ. Learning as a function of time. *J Educ Res*. 1980;73(4):183-194.
14. Schmidt HG, Rotgans JI, Yew EHJ. The process of problem-based learning: what works and why. *Med Educ*. 2011;45(8):792-806.
15. van Blankenstein FM, Dolmans DHJM, van der Vleuten CPM, Schmidt HG. Which cognitive processes support learning during small-group discussion? The role of providing explanations and listening to others. *Instru Sci*. 2011;39:189-204.
16. Schmidt HG, De Volder M, De Grave 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 Psychol*. 1989;81(4):610-619.

17. Loyens SMM, Jones SH, Mikkers J, van Gog T. Problem-based learning as a facilitator of conceptual change. *Learn Instr*. 2015;38:34-42.
18. VanLehn K, Jones RM, Chi MTH. A model of the self-explanation effect. *J Learn Sci*. 1992;2(1):1-59.
19. Chi MTH, de Leeuw N, Chiu M-H, LaVancher C. Eliciting self-explanations improves understanding. *Cognitive Sci*. 1994;18(3):439-477.
20. Rotgans JI, Schmidt HG. The role of teachers in facilitating situational interest in an active-learning classroom. *Teaching Teach Educ*. 2011;27(1):37-42.
21. Rotgans JI, Schmidt HG. Interest development: Arousing situational interest affects the growth trajectory of individual interest. *Contemp Educ Psychol*. 2017;49:175-184.
22. 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 Instr*. 2015;39:72-87.
23. 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: North-Holland Publishing Co; 1975:1-249.
24. Schiefele U, Krapp A. Topic interest and free recall of expository text. *Learn Indiv Diff*. 1996;8(2):141-160.
25. Fraenkel JR, Wallen NE. Content analysis. In: Fraenkel JR, Wallen NE, eds. *How to Design and Evaluate Research in Education*, 7th edn. New York: McGraw-Hill; 2009:472-498.
26. Field A. Exploring data. In: Field A, ed. *Discovering statistics using IBM SPSS*, 3rd edn. Los Angeles: Sage; 2009:85-124.
27. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*. 2013;4:863.
28. Ormrod JE. Memory and its components. In: Ormrod JE, ed. *Human Laearning*, 7th edn. Boston: Pearson; 2016:178-205.
29. Larsen DP, Buttler AC, Roediger HL 3rd. Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med Educ*. 2013;47(7):674-682.
30. Byrne BM. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. New York: Taylor & Francis Group; 2010.
31. 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(8):796-805.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ribeiro LMC, Mamede S, de Brito EM, Moura AS, de Faria RMD, Schmidt HG. Exploring mechanisms underlying learning from deliberate reflection: An experimental study. *Med Educ*. 2021;55:404–412. https://doi.org/10.1111/medu.14410