Title	How cognitive engagement fluctuates during a team-based learning session and how it predicts academic achievement
Author(s)	Rotgans, Jerome Ingmar; Schmidt, Hendricus Gerard; Rajalingam, Preman; Wong, Joey Ying Hao; Canning, Claire Ann; Ferenczi, Michael Alan; Low-Beer, Naomi
Citation	Rotgans, J. I., Schmidt, H. G., Rajalingam, P., Wong, J. Y. H., Canning, C. A., Ferenczi, M. A., et al. How cognitive engagement fluctuates during a team-based learning session and how it predicts academic achievement. Advances in Health Sciences Education,23(2), 339-351.
Date	2017
URL	http://hdl.handle.net/10220/44286
Rights	© 2017 Springer Science+Business Media B.V. This is the author created version of a work that has been peer reviewed and accepted for publication by Advances in Health Sciences Education, Springer Science+Business Media B.V. It incorporates referee’s comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [http://dx.doi.org/10.1007/s10459-017-9801-2].

How Cognitive Engagement Fluctuates During a Team-Based Learning Session and How it Predicts Academic Achievement

Jerome I Rotgans, Ph.D.*

Nanyang Technological University, Singapore

Henk G. Schmidt, Ph.D.

Erasmus University Rotterdam, the Netherlands

Preman Rajalingam, Ph.D.

Nanyang Technological University, Singapore

Joey Wong Ying Hao

Nanyang Technological University, Singapore

Claire Ann Canning, Ph.D.

Nanyang Technological University, Singapore

Michael Alan Ferenczi, Ph.D.

Nanyang Technological University, Singapore

Naomi Low-Beer, M.D.

Nanyang Technological University, Singapore

*Corresponding author: Jerome I Rotgans, Nanyang Technological University, Lee Kong Chian School of Medicine, 11 Mandalay Road, Singapore 308232, RotgansResearch@gmail.com, (+65) 9172 5213

Abstract

The objective of the paper is to report findings of two studies that attempted to find answers to the following questions: (1) What are the levels of cognitive engagement in TBL? (2) Are there differences between students who were more exposed to TBL than students who were less exposed to TBL? (3) To which extent does cognitive engagement fluctuate as a function of the different activities involved in TBL? And (4) How do cognitive engagement scores collected over time correlate with each other and with academic achievement? The studies were conducted with Year-1 and -2 medical students enrolled in a TBL curriculum (N=175, 62 female). In both studies, six measurements of cognitive engagement were taken during the distinct TBL activities (preparation phase, individual/team readiness assurance test, burning questions, and application exercises). Data were analysed by means of one-way repeated-measures ANOVAs and path modelling. The results of the repeated-measures ANOVA revealed that cognitive engagement systematically fluctuated as a function of the distinct TBL activities. In addition, Year-1 students reported significantly higher levels of cognitive engagement compared to Year-2 students. Finally, cognitive engagement was a significant predictor of performance (β = .35). The studies presented in this paper are a first attempt to relate the different activities undertaken in TBL with the extent to which they arouse cognitive engagement with the task at hand. Implications of these findings for TBL are discussed.

Keywords: Academic achievement; cognitive engagement; medical education; micro-analytical measurement; team-based learning.

Introduction

Cognitive engagement is a psychological construct indicating the attempt of a student to deeply understand subject matter he or she is supposed to learn. It is mostly defined as a stable characteristic of students—some students tend to be more cognitively engaged in their education than others, irrespective of the task at hand—although most researchers believe that it is to some extent influenced by other motivational variables such as goal orientation (Greene & Miller, 1996; Meece, Blumenfeld, & Hoyle, 1988). More recently however, interest is growing with regard to situationally-determined cognitive engagement. That is, cognitive engagement that emerges in response to a particular task (Chlapana, 2016; Rotgans & Schmidt, 2011; Schmidt, Benzing, & Kamer, 2016). The purpose of the present article is to report on the emergence of cognitive engagement in response to the various activities that make up the team-based learning classroom and to relate cognitive engagement to performance in those classrooms.

Team-based learning (TBL) is an approach to education that combines elements of direct instruction (prescription of subject matter to be studied, large class lecturing) with more open, small-group-based assignments (Michaelsen, Knight, & Fink, 2002). According to Michaelsen and Sweet (2008) it is not enough for a teacher to simply transmit content. Students should also be enabled to *apply* what they have learned to ensure long-term retrieval of knowledge acquired and its future use. A TBL-cycle usually consists of three phases: (1) A preparation phase; (2) A readiness assurance phase; and (3) An application phase (Parmelee & Michaelsen, 2010). The preparation phase consists of individual learning activities, undertaken outside the classroom, at home or in the library. In addition to books and articles, preparation materials may include recorded lectures and other digital resources. The readiness assurance phase is intended to diagnose whether students have acquired sufficient understanding of the subject studied and to remediate possible shortcomings. First, students engage in an individual closed-book knowledge test (iRAT or Individual Readiness Assurance Test) to assess the extent to which the learning objectives have been attained. This typically consists of 15 to 25 multiple choice questions. Second, the same test is repeated by the students in small teams of 5-6 students (tRAT or Team Readiness Assurance Test). During the tRAT students discuss and come to a consensus on their team answers, and the correct answers are then revealed. This form of peer teaching enables misconceptions to be clarified and gaps to be filled. Often, the small-group discussion leads to additional questions about the subject-matter, sometimes referred to as "burning questions." These questions are dealt with by an instructor who gives feedback, and may provide further clarification by giving a short lecture. Phase 2 usually takes about an hour of class time. Phase 3 consists of application-oriented activities. Here students are confronted with problems, assignments, and exercises encouraging them to apply what they have learned in the previous phases. This phase is characterised by the "4S" principles (Haidet et al., 2012): all student teams work on the *same* problem, which must be *significant*. Furthermore, student teams are required to make a *specific choice* from a limited list of options and report their responses *simultaneously*. Students engage in these activities largely (but not exclusively) through small-group discussion accompanied by teacher feedback. This latter part of TBL may last for 1 to 4 hours.

Studies that investigated the effects on TBL on student learning and achievement are rare. Most articles confine themselves to descriptive analyses of how TBL is conducted and what the pitfalls are with regard to this relatively new approach to post-secondary education (Burgess, McGregor, & Mellis, 2014; Davidson, 2009; Parmelee & Michaelsen, 2010). Only a few present data, mostly on student satisfaction (Vasan, DeFouw, & Compton, 2009). We were able to find a few more studies in which achievement data of TBL classrooms, relative to more conventional forms of education were reported (Fatmi, Hartling, Hillier, Campbell,

& Oswald, 2013; Koles, Stolfi, Borges, Nelson, & Parmelee, 2010; McInerney & Fink, 2003; Nieder, Parmelee, Stolfi, & Hudes, 2005). There is some evidence that TBL has a positive effect on learning, in particular among weaker students (Koles et al., 2010). Another study found that TBL with challenging projects improved the students' comprehension and retention of information, critical thinking, and attitudes about the course (McInerney & Fink, 2003). Even less is known about the mechanism that mediates between TBL and student achievement. Many authors suggest that that TBL encourages deeper engagement with the subject at hand, which would translate into better performance (Clarke & Wells, 2008; Haidet et al., 2012; Kelly et al., 2005). There is however, presently no study demonstrating that TBL increases cognitive engagement and that cognitive engagement in turn influences achievement. The present study sought to establish these relationships.

Cognitive engagement in the classroom can be characterised as a psychological state in which students put in a significant amount of effort to gain understanding of a topic at hand (Corno & Mandinach, 1983) and in which students persist studying over a prolonged period of time (Rotgans & Schmidt, 2011). Cognitive engagement has traditionally been operationalised by measuring students' homework completion, participation in extra-curricular activities, class attendance, and their general interactions with teachers and peers (Appleton, Christenson, Kim, & Reschly, 2006). This description of cognitive engagement suggests that it is considered by most authors a more or less stable trait, independent of its context. In a previous contribution, we have suggested that cognitive engagement is more or less dependent on the task at hand because the task determines to which extent students are willing to invest time and effort. For instance, working in small teams and engaging in discussions, preparing for class by engaging in self-study, or listening to a lecture is likely to result in different levels of cognitive engagement (Rotgans & Schmidt, 2011).

Assuming that cognitive engagement may vary over time as a function of the particular activities undertaken by a student, it is not sufficient to measure the construct once, for instance at the end of a course, as it is typically done in the older cognitive engagement research ("Overall, how engaged were you with the topic at hand?"). Ideally, students should be continuously monitored in terms of the depth of their engagement throughout a learning event. This is not possible yet. We therefore decided to briefly interrupt students several times while they were being engaged in learning and required them to respond to a four-item rating scale, that took them less than a minute to complete. These measurements coincided with the different phases of TBL (see Figure 1).

Insert Figure 1 about here

This research design enabled us to address the following research questions: (1) What are the levels of cognitive engagement in TBL? (2) Are there differences between students who were more exposed to TBL than students who were less exposed to TBL? (3) To which extent does cognitive engagement fluctuate as a function of these activities? And (4) How do cognitive engagement scores collected over time correlate with each other and with academic achievement? We hypothesised that cognitive engagement should be generally relatively high in TBL, considering its reputation to actively engage students in learning. In addition, we hypothesised that students who were less exposed to TBL tend to be more cognitively engaged since this instructional approach is novel to them and constitutes an exciting alternative to conventional teaching. This excitement, however, is expected to wear off to some extent with prolonged exposure. Second, we expected to observe phase-dependent changes during TBL (e.g., cognitive engagement was expected to be lower during the preparation phase as compared to application phase during which students try to apply what they have learned). Lastly, we expected to find a positive correlation between each measure of cognitive engagement and academic achievement (more engagement should result in higher academic achievement).

To test these hypotheses two studies were conducted. In Study 1 a group of first-year medical students engaged in a TBL exercise. Over the course of a day, their cognitive engagement was measured six times. We will further refer to these focused repeated measurements as a "micro-analytical measurement approach." In Study 2, which was a replication and extension of Study 1, second-year medical students were involved. Their cognitive engagement was measured six times which was then related to achievement using path analysis.

Method

Educational Context

At the institution this study was conducted, TBL is the main mode of instruction during the first two years of medical training. During these two years, students are exposed to two TBL sessions per week. All students are required to attend an induction programme that prepares them for TBL. The first-year students that participated in Study 1 were exposed to three TBL sessions prior to the study. The second-year medical students participating in Study 2 had just completed one year of TBL. The TBL sessions were facilitated by two faculty members, one a TBL process expert (the facilitator), and one a subject-matter expert. Students were assigned to teams of six. Prior to coming to class students were assigned pre-reading materials about the topic (Study 1: Amino Acids, Proteins, Membranes and Organelles – Study 2: Fever in childhood, including pathogenic mechanisms, infections, and multi-system physiological response to sepsis). Once in class each student completed 25 multiple-choice questions individually on their electronic tablet (iRAT), and then completed the same questions collaboratively with their team (tRAT). The iRAT and tRAT were closed-book activities. During the tRAT the teams receive automated electronic feedback on the accuracy of their answer after responding to each question. After completing the tRAT, teams wrote down specific questions which they would like further clarification on. These questions were electronically captured and prioritised by faculty. The questions were then reassigned to student teams to prepare a response by consulting available resources, after which a facilitated class discussion was held. Subject matter experts provided immediate corrective and confirmatory feedback during the discussion. This process is referred to as *burning questions*. The final phase of TBL was the application exercise where students, in their teams, were presented with clinical or scientific problems they had to solve. Each of the teams had to work on the same problem and report answers simultaneously. Individuals were then called upon to explain their team's answers to the rest of the class, leading to a dialogue with peers and the subject matter expert. See for a visual overview of the TBL phases Figure 1.

Study Design

Both studies were identical in the sense that repeated measurements of cognitive engagement were taken during each phase of TBL. Six in total; the first when students came to class to gauge their level of understanding during preparation; the second during the iRAT; the third during tRAT; the fourth during generation of burning questions; the fifth during team discussion of the application exercises; and the sixth during class discussion of the application exercises. See Figure 1 for an overview. This enabled measurement of cognitive engagement at critical moments during TBL and plotting of a graph that could reveal whether there was a pattern in the propagation of cognitive engagement over time. Although, we expected to find some mean-level differences between first-year and second-year students in terms of their mean level of cognitive engagement, we expected to observe a similar pattern in how cognitive engagement fluctuated as a function the respective TBL activities. Only for the second Study a measure of academic achievement was administered at the end of the TBL session to explore how cognitive engagement predicted achievement.

Participants

For Study 1, 106 first-year undergraduate medical students from the Lee Kong Chian School of Medicine Singapore participated. Their average age ranged from 18-21, with a mean of 19 years (*SD* = .93). There were 70 males and 36 females.

For Study 2, 69 second-year undergraduate medical students from the Lee Kong Chian School of Medicine Singapore participated. Their average age ranged from 20-23, with a mean of 21 years (*SD* = .89). There were 44 males and 26 females.

Participation was voluntary and written consent was obtained prior to the study. The study was approved by the Institutional Review Board of Nanyang Technological University, Singapore.

Materials

Cognitive Engagement. A *Situational Cognitive Engagement Measure* was used in the study (Rotgans & Schmidt, 2011). The instrument consisted of four items: *Presently* (1) I am engaged with the topic at hand; (2) I put in a lot of effort understanding the topic; (3) I wish I could still continue for a while; and (4) I am so involved that I forget everything around me. All items were scored on a 5-point Likert scale, ranging from 1 (*not true at all*) *to* 5 (*very true for me*). Hancock's coefficient *H* was calculated as a reliability measure. The coefficient *H* is considered a more accurate measure of reliability than the much-used Cronbach's alpha (Hancock & Mueller, 2001; Sijtsma, 2009). Its recommended cut-off value is .70. The coefficient *H* for Study 1 ranged from .72 (measure 1) to .88 (measure 3) with an average value of .81. For Study 2 it ranged from .71 (measure 1) to .90 (measure 4) with an average value of .81. These values are indicative of adequate reliability of the situational cognitive engagement measure.

Academic Achievement. Only for Study 2, a knowledge test with open-ended items was administered to measure students' academic achievement for the TBL session. The open-ended items were derived from the learning objectives for that session (e.g., recall the pathogenic mechanisms of fever). There were 18 open-ended, short-answer items (e.g., give an example of an endogenous pyrogen). The responses were scored by two independent raters using a scoring rubric. One mark was assigned to a correct answer (maximum score was 18). The agreement between raters was .99. Inconsistencies were resolved through discussion.

Procedure

The situational cognitive engagement questionnaire was administered six times during the TBL session using the online platform Qualtrics (2017). Students responded by means of their tablet computers. To adequately measure students' cognitive engagement during the TBL session, we operationalised the study in such a manner that half-way during a TBL activity (e.g., iRAT) they were asked to respond to the four-item survey. See Figure 1 for a schematic overview of the TBL activities and measurement points. This interruption took less than a minute and enabled us to measure their cognitive engagement in situ. This was, however, not possible during the preparatory phase where students were typically off campus and prepared for their lesson in their own time. Therefore, the first administration of the situational cognitive engagement measure took place at the very beginning of the TBL session and was a retrospective measure to determine how cognitively engaged students were during the preparatory phase. The next administrations were approximately half-way during the iRAT, the tRAT, the burning questions class discussion, during the application exercise when the teams discuss the questions, and finally when students reported their application exercise answers in class. Based on previous research, we considered scores on the cognitive engagement scale as low if they were 3 (out of 5), moderate if they were between 3 and 4, and high if they were higher than 4 (Rotgans & Schmidt, 2011). In Year 2, at the conclusion of the session, the knowledge test was administered.

Analysis

For Study 1, we computed mean scores and generated a graph to examine the emergence of cognitive engagement over time. Second, a one-way within-subjects ANOVA was conducted to examine fluctuations in participants' cognitive engagement during the TBL session.

For Study 2, we followed the same procedure as with Study 1 by generating mean values, a graph and conducting a one-way within-subjects ANOVA. We were then able to compare the results between both data set and examined if we succeeded in replicating the findings involving different students and a different topic. We reasoned that if we could replicate the results, it would add to the generalisability of our findings. The analyses were conducted by using the statistical package IBM SPSS (version 24).

In the Year-2 group, as a second step, we tested a path model, linking each measurement of cognitive engagement and the last measure of cognitive engagement predicting academic achievement. To examine the goodness-of-fit for the path model, we generated the Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), and Comparative Fit Index (CFI) along with the χ^2 statistic. Cut-off values of .06 (RMSEA), .09 (SRMR) and .95 (CFI) were used in the analysis (Hu & Bentler, 1999). The analysis was conducted using IBM SPSS AMOS (version 23).

Results

Study 1: Pattern of cognitive engagement

The results of the one-way within-subjects ANOVA yielded a significant time effect Wilk's Λ = .44, $F(5, 525) = 25.87$, $p < .001$, $\eta^2 = .56$. Planned pair-wise comparisons revealed that all differences in mean scores between the subsequent cognitive engagement measures were statistically significant ($p < .05$). The data suggest that cognitive engagement was significantly lower during the preparation phases ($M = 3.51$). It significantly increased during iRAT ($M = 3.95$) and peaked during tRAT ($M = 4.13$). A significant decrease was observed between tRAT and the burning questions ($M = 3.74$), when the content expert answered students' burning questions. Subsequently, cognitive engagement increased significantly when students discussed the application exercises within their teams ($M = 4.01$). Finally, during class discussion of the application exercises, cognitive engagement again decreased significantly ($M = 3.88$). See Figure 2 for a visual overview of the pattern of cognitive engagement over the six measurement occasions.

Insert Figure 2 about here

See Table 1 for descriptive statistics and zero-order correlations between the cognitive engagement measurements.

Insert Table 1 about here

Study 2: Cognitive engagement leading to academic achievement

The results of the one-way within-subjects ANOVA yielded a significant time effect Wilk's $\Lambda = .50$, $F(5, 340) = 13.03$, $p < .001$, $\eta^2 = .50$. Planned pair-wise comparisons revealed that there were statistically significant differences in mean scores between three pairs of cognitive engagement measures. Similar to the results observed for Study 1, there was a significant increase in cognitive engagement between preparation ($M = 3.12$) and iRAT ($M = 3.30$). Although the increase between iRAT ($M = 3.30$) and tRAT ($M = 3.36$) did not reach statistical significance, the decrease between tRAT and burning questions ($M = 2.81$) did, as did the increase between burning questions and the team discussion of the application exercise (M = 3.55). Finally, the decrease between the cognitive engagement during team discussion of the application exercise and the class discussion of the application exercise was not significant. See Figure 3 for an overview of the pattern of cognitive engagement over the six measurement occasions.

Insert Figure 3 about here

See Table 2 for descriptive statistics and zero-order correlations between the cognitive engagement measurements, the mean score of cognitive engagement, and the knowledge measure.

Insert Table 2 about here

As a next step in in the analysis, we explored the relation between cognitive engagement and the academic achievement measure. Closer examination of the raw correlation coefficients suggests that there is a significant association between the mean cognitive engagement score (CE_{MEAN}) and the knowledge measure, $r = .30$, $p = .011$.

To further explore the relation between cognitive engagement and academic achievement, a path model was tested. In an initial model, we allowed each measure of cognitive engagement to regress on the next measure and on academic achievement. The results of the path analysis revealed that the data did not fit this model well: $\chi^2(10) = 73.41$, $p < .001$; CFI = .80; RMSEA = .31 (90% CI: .24-.37); SRMR = .21. Consulting the Modification Indexes suggested two alterations to the model. First, the cognitive engagement measure pertaining to the burning questions did not fit in the sequence of cognitive engagement measures; a better fitting model would be obtained if this measure would be treated as an offshoot of the sequence rather than being a mediator in it. Second, only the cognitive engagement measurement during the tRAT was a significant predictor of academic achievement. See Figure 4 for the corrected model. The data fitted this model well: $\chi^2(15) = 17.58$, $p = .29$; CFI = .99; RMSEA = .05 (90% CI: .00-.13); SRMR = .05 and resulted in a significantly better model fit as compared with the initial model.

Insert Figure 4 about here

Discussion

The objective of the present study was to find answers to the following questions: (1) What is the level of cognitive engagement in TBL? (2) Are there differences between students who were more exposed to TBL than students who were less exposed to TBL? (3) To which extent does cognitive engagement fluctuate per TBL phase? And (4) How do cognitive engagement scores correlate with each other and academic achievement? To that end, we followed two groups of medical students (Year 1 and 2) during a day in which they were engaged in TBL activities and applied a micro-analytical measurement approach to repeatedly measure their cognitive engagement. We briefly interrupted their work at five points in time, chosen such that they were in the middle of a particular TBL activity. Only the first measure of cognitive engagement was a retrospective measure in which we asked them to look back at the previous day, when they prepared themselves for the session. For Year-2 students (Study 2) we conducted a path analysis in which the six measures of cognitive engagement were related to each other and it was then explored how cognitive engagement predicts academic achievement.

Our first finding is that cognitive engagement was moderate to high during the various TBL activities, lending credence to the idea that TBL indeed fosters cognitive engagement in students. Second, cognitive engagement differed significantly between Year-1 (Study 1) and Year-2 (Study 2) students ($F(1, 173) = 57.51$, $p < .001$, $\eta^2 = .25$). The data suggest that cognitive engagement measurement for Year-1 students (Study 1: $M_{CE} = 3.87$, $SD = .48$) was significantly higher than cognitive engagement for Year-2 students (Study 2: $M_{CE} = 3.27$, $SD = .56$). This outcome suggests that TBL is perhaps more engaging for students for whom TBL is a novel alternative to conventional education as was the case for our Year-1 students of Study 1. These students just entered medical school and their previous student career was largely dependent on conventional teaching. This explanation is in line with research that suggests that the excitement of educational innovations, such as using tablet computer in school, wear off over time (Cumming & Rodriguez, 2013). A competing explanation would be that the nature of the topic discussed during the TBL sessions resulted in the observed differences. The Year-1 topic was simply more engaging than the Year-2 topic. This is however unlikely. Medical students prefer clinical topics over basis-science topics. If the topic of the sessions would be important, then the results would likely be the reverse of what we have actually found.

A third observation is that there is quite some variation in cognitive engagement during the day. Many changes were statistically significant, which suggests that cognitive engagement does not seem to be a static construct, but something that is sensitive to different educational activities. For instance, in both studies, cognitive engagement increased significantly between preparation and iRAT, with the preparation being relatively less engaging. This increase in cognitive engagement may be due to the fact that since students prepared for the session, time had passed (24-48hrs), which is sufficient time to allow for memory consolidation to occur (Lee, 2008; Nadel & Moscovitch, 1997; Stickgold, 2005).

Knowledge consolidation may have enabled the students to better recall what was learned previously and manifested itself through their heightened cognitive engagement during the iRAT. However, note that cognitive engagement was measured retrospectively and not during actual self-study a day prior to the TBL session. This is a potential confounder that needs to be taken under consideration.

In addition to the above observation, there is some evidence in our data to suggest that students are significantly more engaged when they work together with peers in small groups which is the case during the tRAT and the first part of the application exercise. At these points in time, cognitive engagement was consistently rated higher. When class activities are scheduled in which the teacher had the lead, such as the burning questions phase and class discussion of the application exercise, cognitive engagement was consistently lower. This pattern could be replicated across both years. This point is supported by research suggesting that cognitive engagement is positively influenced by feeling autonomous from the direct influence of a teacher and when students take the lead in their learning (Appleton et al., 2006; Assor, Kaplan, & Roth, 2002; Greene, Miller, Crowson, Duke, & Akey, 2004; Rotgans & Schmidt, 2011).

Examining the path model and correlations between cognitive engagement measurements, it is interesting to note that although cognitive engagement fluctuates over time, the correlations between adjacent measures are high (and fade as a function of distance in time). This suggests that there is perhaps a trait component visible in our measure: students who display higher levels of cognitive engagement on one measure also demonstrate higher cognitive engagement on subsequent measures.

Examining the path model more closely, there seem to be two irregularities pertaining to cognitive engagement measure 3. First, inspecting the Modification Indexes suggested that cognitive engagement 4 during the burning questions does not fit well in the sequence of the cognitive engagement measures. Thus, we allowed it to branch off. The reason that this measure of cognitive engagement did not fit in the sequence may be due to the fact that only two teams (out of 12) submitted one burning question each. As a result, there was no lengthy and engaging discussion between the teacher and students. To test this explanation, we generated a path model with the data from Study 1. We knew that during that session, substantially more burning questions were raised (30 in total), with each team contributing at least one. Testing a path model with these data suggest that a significantly better model fit can be obtained if the cognitive engagement measure is included in the sequence and does not branch off. Overall, our data imply that only if a sufficient number of burning questions are generated and raised in class, cognitive engagement is part of the main sequence.

The second abnormality is that cognitive engagement measure 3 was a predictor for academic achievement and not, as one may expect, the final measure of cognitive engagement at the end of the TBL sequence. Only the path coefficient between cognitive engagement measure 3 and academic achievement reached statistical significance ($\beta = .35$, $p < .01$), explaining about 12% of the variance in the academic achievement. This outcome makes sense if one knows that academic achievement was operationalised by a test that was derived from the learning objectives stated for this TBL session. At the institution this study was conducted, it is intended that the learning objectives only cover what students are supposed to learn up to the readiness assurance phase (i.e., tRAT and burning questions). Thus, what comes thereafter during the application exercise is not captured by the learning objectives and consequently also not by the test we devised. With this in mind, it makes sense that only a significant association could be found between cognitive engagement 3 at the end of the readiness assurance phase and academic achievement.

The study has some limitations. Using questionnaires to measure cognitive constructs such as engagement in learning is at best an approximation. It would have been better to use one of the methods from the cognitive psychology laboratory, such as measuring the psychological refractory period, the amount of time it takes to switch from one stimulus to another, signifying depth of attention. However, such methods are hardly amenable to use in classroom settings. Using a reliable short questionnaire that can be taken more than once, is the best option available. Another potential issue is that we interrupted students several times during TBL and that these interruptions may have a negative effect on subsequent measurements. Note, however, that these interruptions were of very short duration (less than a minute) and our data did not suggest that interruption themselves had a negative effect on cognitive engagement in the subsequent interruption. Another limitation of the present study is that our sample consisted of medical students only. In order to draw more generalizable conclusions of our findings, the study needs to be replicated with a larger sample across multiple disciplines and age groups.

In conclusion, the use of the micro-analytical measurement approach in studying cognitive engagement in the TBL classroom shows that it fluctuates depending on instructional activities that students are conducting. Our study also shows that sustained cognitive engagement leads to significant knowledge gains in TBL. The micro-analytical measurement approach may also be useful in studying other important characteristics of learning in the TBL classroom, such as the emergence of interest in subject matter, difficulty level of concepts to be mastered, or motivational variables of interest.

References

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*(5), 427-445.

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology, 72*(2), 261-278.

Burgess, A. W., McGregor, D. M., & Mellis, C. M. (2014). Applying Established Guidelines to Team-Based Learning Programs in Medical Schools: A Systematic Review. *Academic Medicine, 89*(4), 678-688. doi:10.1097/acm.0000000000000162

Chlapana, E. (2016). An intervention programme for enhancing kindergarteners' cognitive engagement and comprehension skills through reading informational texts. *Literacy, 50*(3), 125-132.

Clarke, J., & Wells, M. (2008). *Making Technology Work for you: Why Might Academic Staff Want to Engage With and Promote e-Learning? A Case Study*. NR Reading Academic Conference LTD.

Cumming, T. M., & Rodriguez, C. D. (2013). Integrating the iPad into language arts instruction for students with disabilities: Engagement and perspectives. *Journal of Special Education Technology, 28*(4), 43-52.

Corno, L., & Mandinach, E. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist, 18*(2), 88-108.

Davidson, L. (2009). *Educational innovation in an undergraduate medical course: Implementation of a blended e-learning, team-based learning model.*

Fatmi, M., Hartling, L., Hillier, T., Campbell, S., & Oswald, A. E. (2013). The effectiveness of team-based learning on learning outcomes in health professions education: BEME Guide No. 30. *Medical Teacher, 35*(12), e1608-e1624.

Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology, 21*(2), 181-192.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology, 29*(4), 462-482.

Haidet, P., Levine, R. E., Parmelee, D. X., Crow, S., Kennedy, F., Kelly, P. A., . . . Richards, B. F. (2012). Perspective: Guidelines for Reporting Team-Based Learning Activities in the Medical and Health Sciences Education Literature. *Academic Medicine, 87*(3), 292-299. doi:10.1097/ACM.0b013e318244759e

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent systems. In R. Cudeck, S. d. Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future - A festschrift in honor of Karl Jöreskog* (pp. 195-216). Lincolnwood, IL: Scientific Software International.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1-55.

Kelly, P. A., Haidet, P., Schneider, V., Searle, N., Seidel, C. L., & Richards, B. F. (2005). A Comparison of In-Class Learner Engagement Across Lecture, Problem-Based Learning, and Team Learning Using the STROBE Classroom Observation Tool. *Teaching and Learning in Medicine, 17*(2), 112-118. doi:10.1207/s15328015tlm1702_4

Koles, P. G., Stolfi, A., Borges, N. J., Nelson, S., & Parmelee, D. X. (2010). The Impact of Team-Based Learning on Medical Students' Academic Performance. *Academic Medicine, 85*(11), 1739-1745. doi:10.1097/ACM.0b013e3181f52bed

Lee, J. L. (2008). Memory reconsolidation mediates the strengthening of memories by additional learning. *Nature neuroscience, 11*(11), 1264-1266.

McInerney, M. J., & Fink, L. D. (2003). Team-based learning enhances long-term retention and critical thinking in an undergraduate microbial physiology course. *Journal of Microbiology & Biology Education, 4*(1), 3-12.

Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology, 80*(4), 514-523.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (Eds.). (2002). *Team-based learning: A transformative use of small groups*. Portsmouth. NH: Greenwood publishing group.

Michaelsen, L. K., & Sweet, M. (2008). The essential elements of team-based learning. *New Directions for Teaching and Learning*(116), 7-27.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology, 7*(2), 217-227. doi:10.1016/s0959-4388(97)80010-4

Nieder, G. L., Parmelee, D. X., Stolfi, A., & Hudes, P. D. (2005). Team-based learning in a medical gross anatomy and embryology course. *Clinical Anatomy, 18*(1), 56-63. doi:10.1002/ca.20040

Parmelee, D. X., & Michaelsen, L. K. (2010). Twelve tips for doing effective Team-Based Learning (TBL). *Medical Teacher, 32*(2), 118-122.

Rotgans, J. I., & Schmidt, H. G. (2011). Cognitive engagement in the problem-based learning classroom. *Advances in Health Sciences Education, 16*(4), 465-479. doi:10.1007/s10459-011-9272-9

Schmidt, M., Benzing, V., & Kamer, M. (2016). Classroom-based physical activity breaks and children's attention: cognitive engagement works! *Frontiers in Psychology, 7, Article 1474*.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's Alpha. *Psychometrika, 74*(1), 107-120.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437*(7063), 1272-1278. doi:10.1038/nature04286

Vasan, N. S., DeFouw, D. O., & Compton, S. (2009). A Survey of Student Perceptions of Team-Based Learning in Anatomy Curriculum: Favorable Views Unrelated to Grades. *Anatomical Sciences Education, 2*(4), 150-155. doi:10.1002/ase.91

Table 1

Study 1: Mean Scores of the Cognitive Engagement (CE) Measurements and Zero-Order Correlations Between the Measurements (Year 1, N = 106)

Variable	CE 1	CE 2	CE 3	CE 4	CE 5	CE 6
CE 1	1	.51**	.42**	.45**	.40**	.30**
CE 2		1	.73**	.55**	.41**	.36**
CE 3			1	.59**	.52**	.42**
CE 4				1	.66**	.61**
CE 5					1	.67**
CE 6						1
Mean	3.51	3.95	4.13	3.74	4.01	3.88
(SD)	(.52)	(.60)	(.64)	(.67)	(.65)	(.65)

Note: ** $p < .01$ level and * $p < .05$ level.

Table 2

Study 2: Mean Scores of the Cognitive Engagement (CE) Measurements, the Mean Cognitive Engagement Score (CE_{MEAN}), and Knowledge and Zero-Order Correlations Between the Measurements. (Year 2, N = 69)

Variable	CE 1	CE 2	CE 3	CE 4	CE 5	CE 6	CE_{MEAN}	Knowledge
SCE 1	1	.63**	.52**	.51**	.50**	.39**	.70**	.09
SCE 2		1	.86**	.53**	.75**	.60**	.88**	.25*
SCE 3			1	.59**	.83**	.72**	.92**	.35**
SCE 4				1	.48**	.47**	.74**	.24*
SCE 5					1	.81**	.89**	.23
SCE 6						1	.81**	.31**
CE_{MEAN}							1	.30*
Knowledge								1
Mean	3.11	3.30	3.36	2.81	3.55	3.47	3.27	15.56
(SD)	(.56)	(.65)	(.70)	(.79)	(.69)	(.69)	(.56)	(1.70)

Note: ** $p < .01$ level and * $p < .05$ level.

Figure 1. Overview of TBL activities and cognitive engagement measurements.

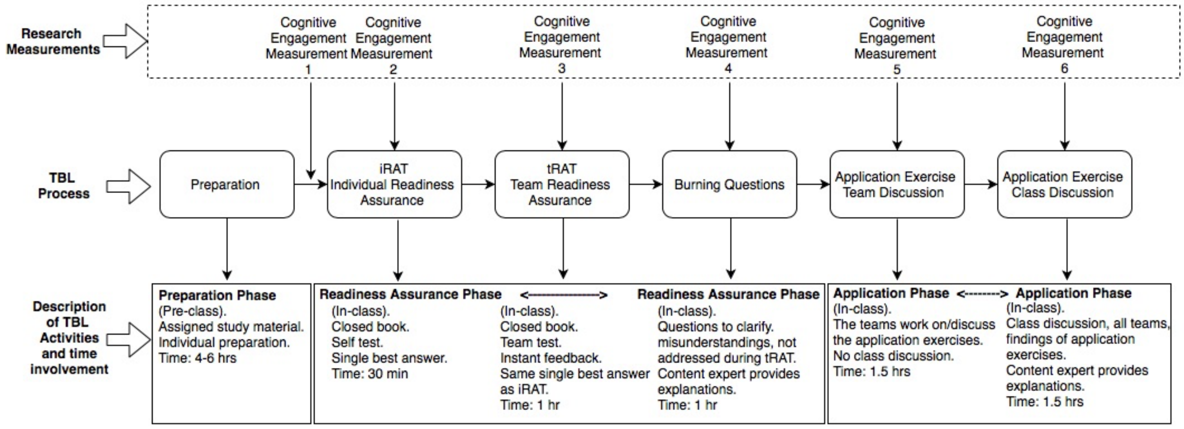
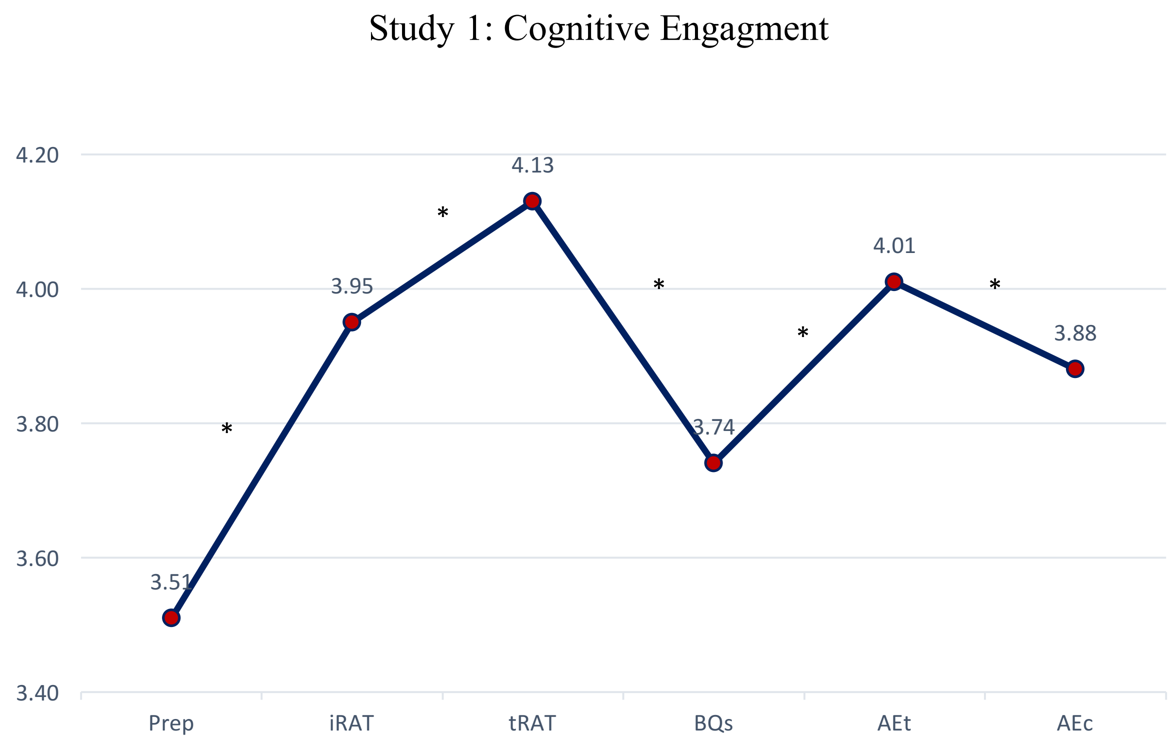
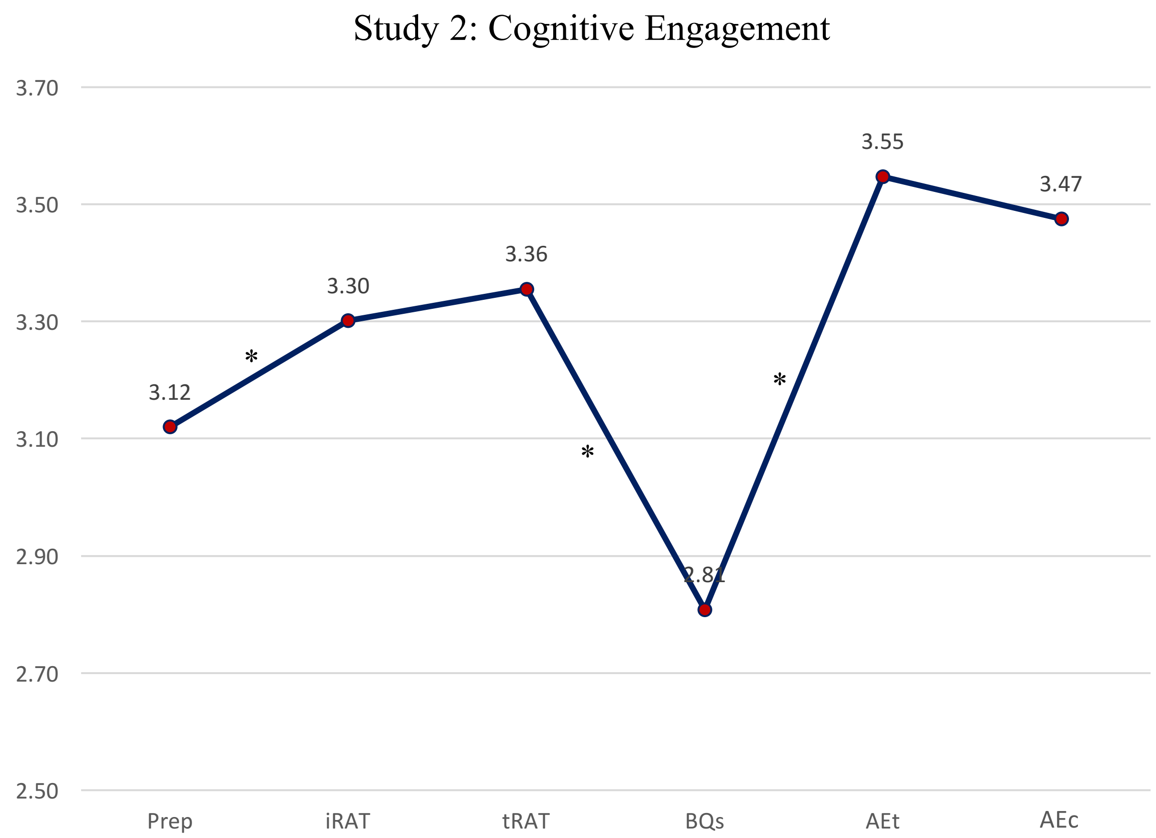


Figure 2. Study 1 (Year 1). Mean levels for cognitive engagement during the TBL session for the preparatory phase (prep), the individual readiness assurance test (iRAT), the team readiness assurance test (tRAT), the class discussion of the burning questions (BQs), the team discussion of the application exercise (AEt) and the class reporting of the application exercise (AEc) (N = 106).



Note: All sequential changes between cognitive engagement measures are significant $p < .05$. Scale from 3.40 to 4.20 on a 5-point Likert scale.

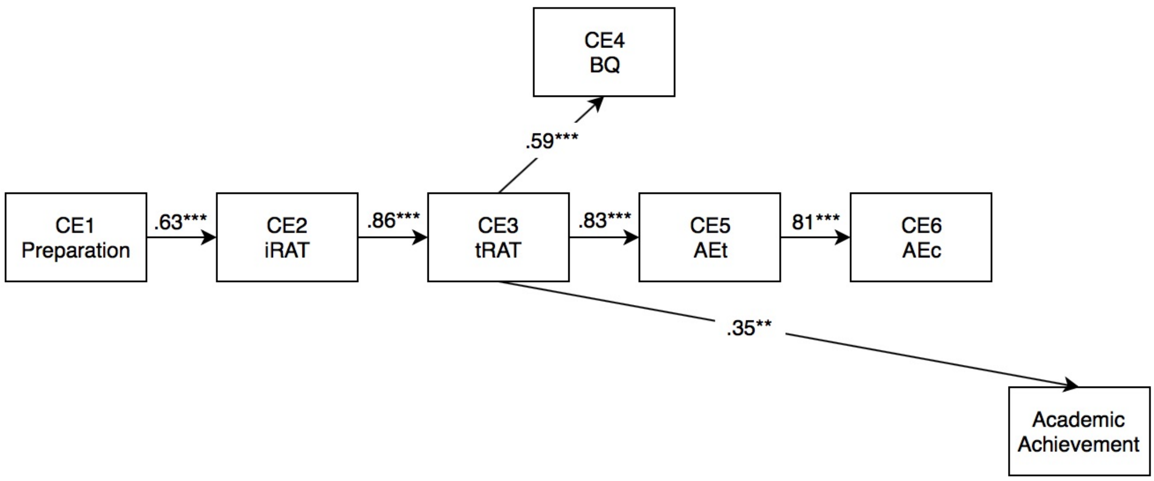
Figure 3. Study 2 (Year 2). Mean levels for cognitive engagement during the TBL session for the preparatory phase (prep), the individual readiness assurance test (iRAT), the team readiness assurance test (tRAT), the class discussion of the burning questions (BQs), the team discussion of the application exercise (AEt) and the class reporting of the application exercise (AEc) (N = 69).



Note: Significant differences between measurements * $p < .05$.

Scale from 2.50 to 3.70 on a 5-point Likert scale.

Figure 4. Study 2. Path model of the cognitive engagement measurements (CE1: cognitive engagement preparation phase, CE2 iRAT: cognitive engagement during the iRAT, CE3 tRAT: the cognitive engagement during the tRAT, CE4 BQ: cognitive engagement during the burning questions, CE5 AEt: cognitive engagement during the team discussion of the application exercise, CE6AEc: cognitive engagement during class reporting of the application exercise) and knowledge test measure (N = 69).



Note: *** $p < .001$, ** $p < .01$, and * $p < .05$ level.