Available online at www.sciencedirect.com
ScienceDirect
Health Professions Education 5 (2019) 294–302
www.elsevier.com/locate/hpe

A Students' Model of Team-based Learning

Jerome I. Rotgans*, Preman Rajalingam, Michael A. Ferenczi, Naomi Low-Beer

Nanyang Technological University, Lee Kong Chian School of Medicine, 11 Mandalay Road, Singapore 308232, Singapore

Received 10 September 2018; received in revised form 23 October 2018; accepted 23 October 2018
Available online 29 October 2018

Abstract

Purpose: Team-based learning (TBL) combines direct instruction with active, collaborative small group learning. This study aimed to elucidate-from the students' perspective-the relations between different elements of TBL. This is expected to provide a better understanding of the inner workings of TBL in education.
Method: Three hundred and thirteen first- and second-year medical students participated in the study. Data about TBL were collected at the end of six teaching blocks, by means of a questionnaire. The data were then combined and subjected to path analysis, which enabled testing of hypothesised relations between three layers of TBL-relevant variables. These were (1) input variables: prior knowledge, teamwork, challenging application exercise, content expert and facilitator; (2) process variables: preparation materials, individual readiness assurance test (iRAT), team readiness assurance test (tRAT); and (3) output variables: learning and topic interest.
Results: Initial analysis resulted in amendments to the hypothesised model. An amended model fitted the data well and explained 43% of the variance in learning and 32% of the variance in topic interest. Content expert had a direct effect on topic interest, as did prior knowledge, teamwork, iRAT and application exercise. Learning was directly influenced by tRAT, application exercise and facilitator, but not content expert.
Discussion: The results of this study demonstrate the inter-relationships of different elements of TBL. The results provide new insights in how TBL works from a students' perspective. Implications of these findings are discussed.

Keywords: Team-based learning; Medical education; Path analysis; Students' model

1. Introduction

Team-based learning (TBL) is a relatively recent development in education, combining elements of both direct instruction and active, collaborative learning in small-groups. TBL is usually structured in three phases.[1] The first phase is the individual preparation phase. During this phase, students study book chapters, articles or digital resources, prescribed by their teachers. Subsequently, the students meet in class to start the next phase. The second phase is the readiness assurance phase. Students first engage in an individual closed-book knowledge test (iRAT or individual readiness assurance test) to check whether their understanding of the learning materials is sufficient. This typically consists of 15 to 25 multiple choice questions. The same test is repeated by the students in small

*Corresponding author.
E-mail address: jerome.rotgans@ntu.edu.sg (J.I. Rotgans).
Peer review under responsibility of AMEEMR: the Association for Medical Education in the Eastern Mediterranean Region

https://doi.org/10.1016/j.hpe.2018.10.003

groups of 5–7 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 may lead to additional questions about the subject-matter, sometimes referred to as "burning questions" for which the teacher provides further clarification. The third phase is the application phase. During this phase, the small groups engage in a series of exercises encouraging them to apply what they have learned. This phase is characterised by the "4 S" principles:[2] 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 through intra-and inter-team discussions, accompanied by teacher feedback.

TBL seems to have a number of advantages over conventional lecture-based education. Since knowledge acquisition is left to the individual student, lectures can be discarded, leaving classroom time free to engage in other activities relevant to learning. The second advantage is that students receive feedback on the extent of their learning, from peers when they discuss answers to the test during tRAT and from the teacher when discussing the burning questions. The third advantage is that students actively engage in the application of what is learned to new problems, assignments, or questions. It is assumed that this attempt to apply what one has learned facilitates its consolidation and later retrieval.[3]

Research studying TBL's effects on student learning is still limited. Most of the papers concentrate on attitudes toward TBL, showing that students tend to prefer TBL over other forms of instruction.[4,5] Others have looked at outcomes, comparing TBL with conventional education. Most of these studies suggest that TBL is more effective in terms of student learning.[6–8] Koles et al.[9] for instance, found an effect on pathology-based questions favouring TBL. McInerney and Fink[10] demonstrated a similar effect on long-term retention and critical thinking in an undergraduate microbial physiology course. Nieder et al.,[11] on the other hand, found no difference in a gross anatomy and embryology course.

Studies aimed at the "inner workings" of TBL are however rare. For instance, one assumption of TBL is that performance during both the tRAT and application exercises build on knowledge acquired during individual preparation (as measured by performance on the iRAT). Is this indeed the case? It is also not clear to which extent the teacher influences the learning process. Since the structured activities in TBL provide a substantial degree of guidance, it is possible that the teacher only plays a minor role when clarifications are needed during the team discussions. In addition: What is the role of collaborative learning in teams? Although TBL is anchored in the premise that learning in teams is beneficial, relatively little is known to which extent the team adds to the learning experience and understanding in TBL. Does the team really play a significant role or is individual self-study during the preparation phase more essential? These are questions to which relatively little empirical answers are available in the TBL literature (for an exception see).[12] We propose that a first step towards advancing our understanding of these issue is to apply a research paradigm that involves students as *observers of their own learning* in TBL using self-report measures.

There is a tendency among educators to interpret student responses to self-report measures as mere expressions of satisfaction. However, students can be intelligent co-researchers of their own learning, because, unlike the professional researcher who often only sees the results of learning in students, they are present at every step. Marsh[13] for instance, has demonstrated that students can be reliable, valid and unbiased judges of the quality of teaching. De Grave and Dolmans[14] showed that the contributions of tutors in problem-based learning can be judged by the students with high inter-rater agreement.

A source of inspiration for the present study was the development and testing of a causal model of problem-based learning that was initiated by Gijselaers and Schmidt[15] and further developed by Van Berkel and Schmidt.[16] These authors distinguish between three types of variables relevant to learning and achievement: *input variables* (level of prior knowledge; quality of the learning resources; quality of teachers), *process variables* (tutorial group activities that support learning; amount of time spent on learning), and *output variables* (student achievement; increased topic interest). Using path modelling, these authors were able to relate the outcomes of learning to process and input variables. For instance, the quality of the problems used in problem-based learning influenced the quality of how the tutorial group functioned, which, in turn, influenced time spent on learning and achievement.[15] Applying these ideas to TBL, we assumed that the three phases of TBL are positively related to each other.[3,17] That is, the preparation phase (preparation materials) influences the readiness assurance phase (iRAT and tRAT), which in turn influences the application phase. Fig. 1 contains a proposed path model, in which influences of input variables on the core process, and their hypothesized effects on output variables are indicated by arrows. The

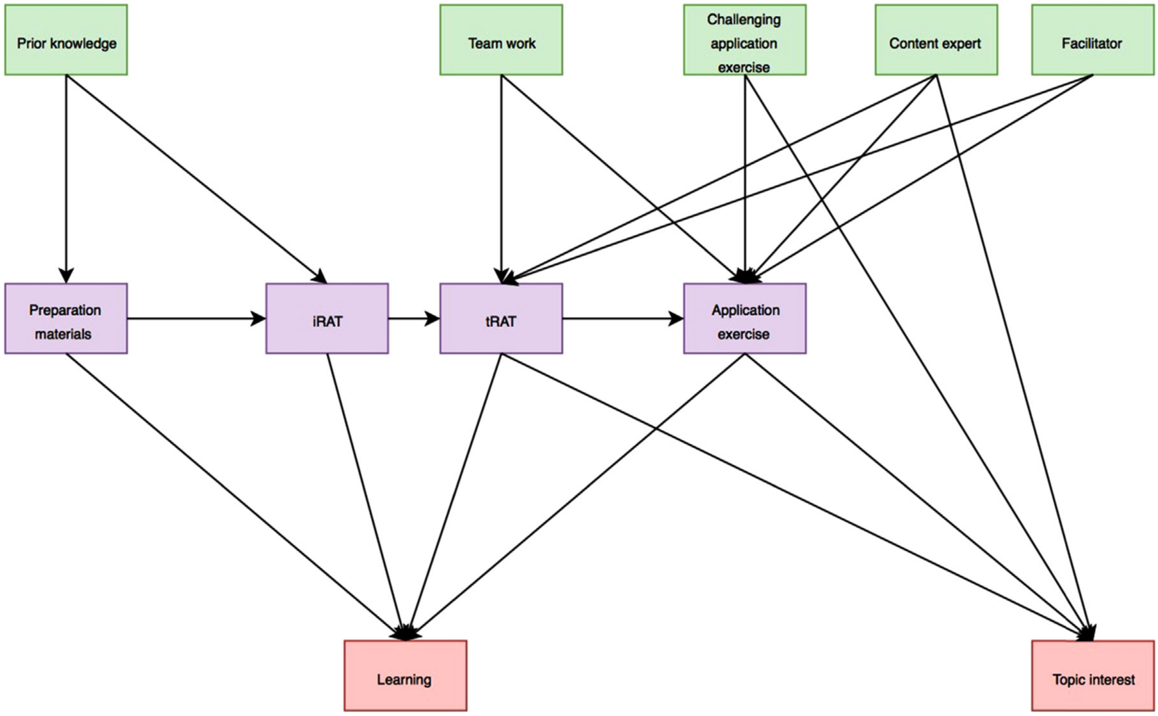

Fig. 1. Theoretical path model of students' perception of TBL with input, process and output variables.

input variables are represented as the top-layer rectangles, whereas the output variables are represented as bottom-layer rectangles.

The path model should be read as follows. The more knowledge the student has, the more will be learned from the preparation materials.[18] The more is learned from these materials, the better the performance on the iRAT, and the greater the potential for subsequent learning. Performance on the iRAT, in turn, positively influences the tRAT, and so on. We will briefly discuss the theoretical and empirical rationale for the hypothesized relationships depicted in Fig. 1. The reader should however bear in mind that the proposed path model is a first attempt to model variables of interest in the context of TBL. It this sense, our study should be considered as exploratory.

In the path model, we further hypothesised that the tRAT and application exercise would only be effective if the teams work well together, hence the linkages between team work and tRAT and application exercise respectively. In our experience of TBL, we noticed that the application exercises are only perceived as effective by the students if they are sufficiently challenging. In other words, if the exercises are too easy and do not stretch a students' knowledge, the application exercises are perceived as not being very effective in helping them learn. To test this hypothesis in the model, we added the input variable "Challenging application exercise."

Furthermore, it should not come as a surprise that the teacher is expected to have a significant influence on the TBL core didactic process as well, in particular, when it comes to providing feedback and explanations during team and class discussions. Thus, we hypothesised that the teacher will have a significant influence on both the tRAT and the application exercise in the model. Note, that in the model in Fig. 1, we distinguished between "Content expert" and "Facilitator" as roles of the teacher. This reflects the educational context in which the present study was conducted, where two teachers with different roles are present throughout the session. For more details, see *Educational Context* in the Method Section.

Finally, the TBL core didactic process is expected to influence prospective outcome variables. The most discernible one is student learning.[19,20] Since potentially all elements of the TBL core process have an effect on student learning, we allowed the preparation materials, the iRAT and tRAT, and application exercise to regress on "Learning" in the model. Besides the acquisition of knowledge, it is conceivable that TBL has other positive outcomes on the students' learning experience. For instance, when TBL was first conceived, one important consideration was to make the topic-to-be-learned more interesting to students.[21] Group discussions, applying what one has learned to real-life contexts are all elements that may stimulate topic interest in students. To explore how TBL results in increased interest in the topic being taught, we included "Topic interest" as an additional outcome variable in the model.

Finally, to test the model described above, we asked first- and second-year medical students of a new medical school in Singapore to respond to a questionnaire containing items aimed at measuring their perception of the various elements of TBL. The ensuing data were subjected to path modelling, under the structural equation modelling paradigm, to test their fit to the (theoretical) model outlined in Fig. 1.

2. Method

2.1. Educational context

This study was conducted at the Lee Kong Chian School of Medicine Singapore (LKCMedicine), where TBL is the main instructional approach for the first two years of a five-year MBBS programme. The TBL process at LKCMedicine adheres to the principles of TBL described in Parmelee et. al and the guidelines put forward by Haidet et. al. TBL is applied consistently in all of the teaching blocks.[22] A distinctive feature about TBL at LKCMedicine is that there is a facilitator and a content expert. Each TBL session is team-taught by both of them, each having distinct areas of responsibility. The content expert is usually a scientist or clinician who developed the learning material and has deep subject matter expertise in the specific topic being learned. In class the main role of the content expert is to guide students in their understanding of content as they engage with the TBL process. The content expert works closely with the TBL facilitator who understands the TBL process and has extensive pedagogical knowledge. In addition to managing the classroom, the facilitator is responsible for questioning students and holding them accountable for their own learning, but is not usually an expert in the subject matter being learned.

2.2. Participants

Three hundred and thirteen first- and second-year undergraduate medical students (36% female and 64% male) from the Lee Kong Chian School of Medicine Singapore participated in the study. The age of the participants ranged from 19 to 24 years, with an average age of 21 years (SD = .92). Responding to the questionnaires used in this study was voluntary and the responses were anonymous. The mean response rate was 53%.

3. Materials

3.1. Team-Based Learning Questionnaire

The items used to measure the TBL learning process were taken from a regular programme evaluation questionnaire that was administered cross-sectionally at the end of each block in 2017. In total, we collected data for six blocks: Introduction to medical sciences, renal and endocrine, cardiorespiratory, gastro-intestinal blood and infection, neuro, ENT & eyes, and reproductive medicine and child health. The following items were used in the analysis. Didactic process variables (TBL core process): (1) "*The TBL preparation materials (narrated PowerPoints, eBooks, face- to-camera lectures etc.) met my learning needs*;" (2) "*The iRAT questions helped me to check my understanding of the important concepts I studied during the pre-reading*;" (3) "*The discussion with my team members during the tRA questions helped me to think critically*;" (4) "*The Application Exercises helped me to understand and/or apply important concepts*." Input variables: (5) "*The topics of the Block were aligned with my prior knowledge*;" (6) "*The team worked well together*;" (7) "*The Application Exercises were sufficiently challenging to promote team discussion*;" (8) "*The sessions were facilitated in ways that helped me learn*;" (9) "*The content experts helped me learn*." Outcome variables: (10) "*The TBL process (pre-reading, iRAT/tRAT, AEs, team and class discussions)* as *a whole helped me to learn important content knowledge and skills*;" and (11) "*The topics we addressed in this block were interesting*." All items were

Table 1
Intra-class Correlation Coefficient for the Six Teaching Blocks.

Block	Responses	ICC	Number of TBL sessions
Year 1: Introduction to medical sciences	40	.81	17
Year 1: Renal and endocrine	45	.72	13
Year 1: Cardiorespiratory	102	.89	17
Year 2: Gastro-Intestinal blood and infection	23	.73	24
Year 2: Neurology, ENT & Eyes	81	.96	15
Year 2: Reproductive medicine and child health	22	.86	17

scored on a 5-point Likert scale, ranging from 1 (*not true at all*) to 5 (*very true for me*).

To establish the reliability of the student judgements measured by the single items stated above, we generated for each of the six blocks an inter-rater agreement score (students being the raters). If there is high agreement between the students on all 11 items, it is indicative that the measurements (and, hence, their observations) are reliable.[23] We used the Intra-Class Correlation coefficient, or ICC, as a measure of agreement.[24] See Table 1 for an overview of the ICC values for each block.

The ICC values were relatively high (> .70), with a mean ICC of .83, suggesting high levels of agreement between students in their observations regarding TBL. This outcome implies that using single items is a reliable approach of quantifying students' observations of TBL.

3.2. Procedure

At the end of each teaching block, Year 1 and 2 students were invited to respond to an end-of-block-evaluation questionnaire. Data were then merged to be used in the path analysis.

3.3. Analysis

Data were analysed by composing a path model. The path model represented three layers of variables; input variables, process variables, and outcome variables. For an overview see Fig. 1. To examine the extent to which our preconceived model fitted the data, 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.[25] The analysis was conducted using AMOS 23.[26] For all analyses, the significance level was set at $p < .05$.

4. Results

As a first step in the analysis, we tested whether the hypothesised model, as depicted in Fig. 1, fits the data. The results of the model fit indices suggest that this was not entirely the case: $\chi^2(25)=170.74$, $p<.001$; CFI=.89; RMSEA=.14 (90% CI: .12–.16); SRMR=.11. See Table 2 for the zero-order correlations and mean values of the variables involved.

A common problem in testing models like these is that the hypothesised model may be too restrictive, which can result in a poorer fitting model.[15] In order to improve the model fit, we inspected the modification indices that were produced with the analysis. The modification indices suggested that a more relaxed model, with more connections, would result in a better model fit.

We made some revisions to the original model and the amended model did indeed fit the data better: $\chi^2(26)=105.85$, $p<.001$; CFI=.94; RMSEA=.099 (90% CI: .08–.12); SRMR=.087. The revised model was a significant improvement over the first model ($\Delta\chi^2=64.89$, $\Delta df=1$, $p<.001$). See Fig. 2 for an overview of the model with its standardised regression weights. The dotted lines highlight the main differences between both models.

The model confirmed that the didactic core process of TBL consists of "Preparation materials" leading to

Table 2
Zero-Order Correlations Between Items of the TBL Questionnaire as Well as Descriptive Statistics (N=313).

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Prior knowledge	1.00										
(2) Team work	.35	1.00									
(3) Challenging application exercise	.30	.40	1.00								
(4) Content expert	.37	.34	.49	1.00							
(5) Facilitator	.39	.35	.50	.49	1.00						
(6) Preparation materials	.49	.22	.24	.17	.32	1.00					
(7) Individual readiness assurance test	.39	.38	.36	.34	.41	.36	1.00				
(8) Team readiness assurance test	.32	.58	.56	.45	.49	.33	.51	1.00			
(9) Application exercise	.33	.37	.63	.52	.46	.28	.36	.52	1.00		
(10) Topic interest	.40	.41	.38	.46	.34	.27	.41	.41	.41	1.00	
(11) Learning	.36	.39	.47	.43	.57	.40	.42	.53	.49	.37	1.00
Mean	3.70	4.34	4.22	4.23	3.93	3.65	4.18	4.14	4.12	4.24	4.09
(SD)	.84	.57	.64	.63	.74	.82	.57	.60	.68	.64	.64

Note: all correlations are significant at the $p<.01$ level (2-tailed).

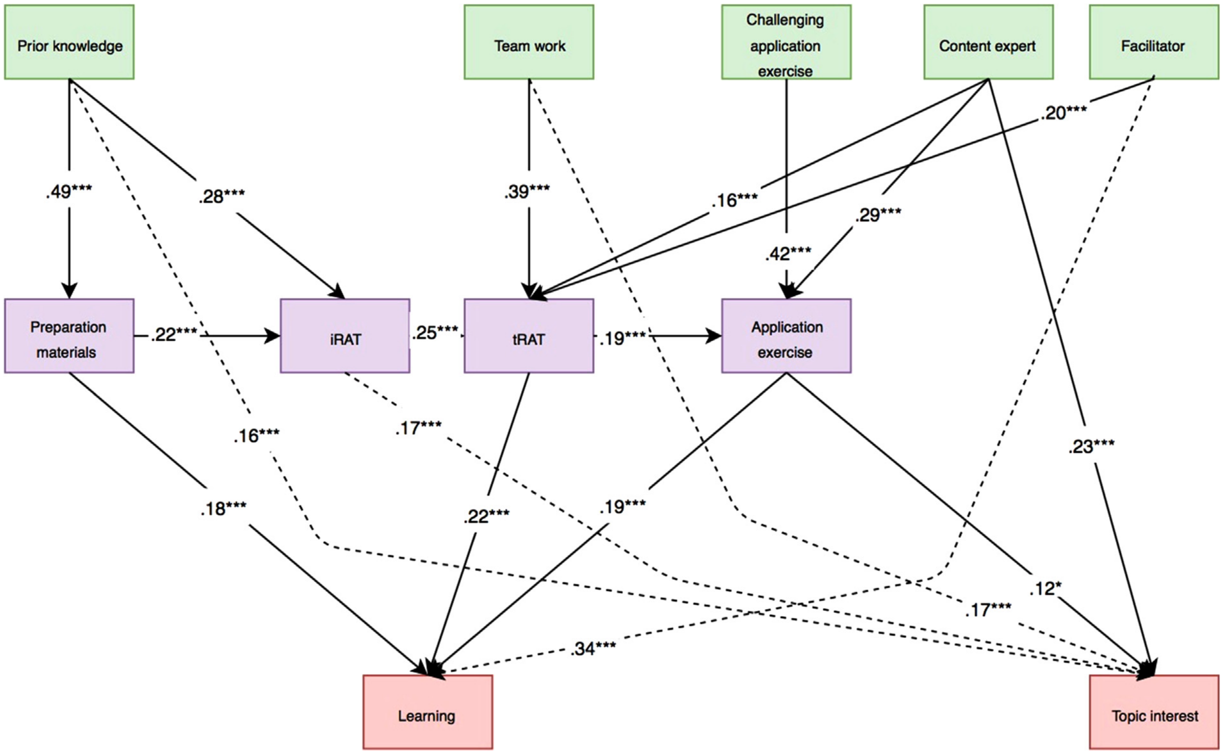

Fig. 2. Modified path model of students' perception of TBL (N = 313) with standardised regression vales. Note: Dotted lines indicate deviations from the theoretical model presented in Fig. 1. *** $p < .001$ level, ** $p < .01$ level and * $p < .05$ level.

"iRAT" (β = .22, $p < .001$), leading to "tRAT" (β = .25, $p < .001$), which in turn leads to the "Application exercise" (β = .19, $p < .001$).

As hypothesised, the input variable "Prior knowledge" was positively associated with "Preparation materials" (β = .49, $p < .001$) and with the "iRAT" (β = .28, $p < .001$). "Team work" had a positive and relatively strong influence on "tRAT" (β = .39, $p < .001$), but was not associated with the application exercise. Instead, it was a predictor of "Topic interest" (β = .17, $p < .01$). In line with our hypothesised model, the more challenging an application exercise was perceived, the more students felt they learned during the application exercise (β = .42, $p < .001$). However, contrary to our predictions, "Challenging application exercise" did not have a direct influence on "Topic interest." This relationship was mediated by "Application exercise" itself (β = .12, $p < .05$).

As predicted in the model, the content expert had a direct influence on the "tRAT" and "Application exercise" (β = .16 and β = .29 $p < .001$ respectively). However, unlike our predictions, the facilitator only had a direct influence on the "tRAT" (β = .20, $p < .001$), but not on the "Application exercise." Instead, the facilitator had a direct influence on learning.

Further, "Learning" was also determined by the quality of the "Preparation materials" (β = .18, $p < .001$) the "tRAT" (β = .22, $p < .001$) and the "Application exercise" (β = .19, $p < .001$), but not by the "iRAT."

The second outcome variable, "Topic interest," was influenced by the "Content expert" (β = .23, $p < .001$) and the "Application exercise" (β = .12, $p < .05$), which was in line with our predictions. However, "Topic interest" was also directly determined by "Prior knowledge" (β = .16, $p < .001$) and the "iRAT" (β = .17, $p < .001$) (and "Team work" as mentioned earlier).

5. Discussion

The purpose of the study reported in this article was to develop insight in the inner workings of TBL. To that end, we asked first- and second-year medical students of a new medical school in Singapore, that uses TBL as its main teaching method, to reflect on their training and to report how the various elements of the TBL arrangement impinge on their learning. Their observations allow for a number of inferences.

First, in the eyes of the students, all elements of the model make sense, it has good face validity. Each element contributes to overall learning and topic interest, either directly, or indirectly through their common relationship with a third variable. For instance, the quality of the content expert does not influence overall learning directly, but indirectly via the application exercises. The different inter-relationships seem "logical."

For instance, the influence of prior knowledge is where one would expect it to be. Interestingly, the temporal sequence of the elements that are considered the "core" of TBL: preparation materials, iRAT, tRAT and application exercise, is neatly replicated in the directionality of the arrows connecting them. Attempts to test the model with reversals of the arrows led to poorer goodness-of-fit.

A second inference is that overall learning is mainly influenced by the tRAT and application exercise. The team and class discussion that occur during these phases of TBL, which are essentially elaborations on and applications of the initial learning, have a direct added value here. This finding provides first support for the general assumption that readiness assurance and application of knowledge in TBL are important ingredients that help students learn.[3,21,27] Unsurprisingly, the preparation materials also had an effect on learning; the better the materials, the more students learn. A surprising finding, however, was that the facilitator, who is primarily responsible for classroom management, had a direct and relatively substantial influence on learning. Since the facilitator's role is primarily classroom management, and not conveying content knowledge, it is currently not clear why the facilitator has such a positive effect on learning. It is possible that in the process of managing the class, the facilitator enacts learning supportive behaviours that have a direct effect on student learning. For instance, the facilitator frequently encourages students to elaborate and generate (burning) questions during the tRAT and calls upon students to respond to their peers' questions. This explains the positive association of the facilitator with the iRAT, but not the direct relation with learning. More research is needed. Overall, the model explained 43% of the variance in learning.

The third inference is that topic interest is not determined by the same variables that determine learning outcomes. Topic interest seems to be influenced by the student him- or herself, the team, and the variables and teacher. Prior knowledge and how much was learned from the iRAT, both person-related variables, had a positive influence on how interesting the topic was perceived. The extent to which the teams worked well together determined how much interest the students experienced in the topic at hand. And third, the content expert, with his or her knowledge of the topic positively influenced topic interest. Overall, the model accounted for 32% of the variance in interest. Finally, note that topic interest is not an input variable but an output variable in the model. Attempts to feed it into the model as an *input* variable were not supported by the empirical data. These findings replicate the problem-based learning literature; increased interest is a function of group functioning and problem quality. In addition, interest is an outcome of learning rather than a starting point.[15,16,28–30]

The students' model of TBL illustrates the unique roles of the educator in this approach to learning and instruction; whether as a facilitator or as a content expert. What stands out is that the context expert has an indirect, as opposed to a direct, influence on learning. This is possibly a consequence of making students individually and in teams more in charge of their own learning. The content expert supports and coaches, but he or she does not lead. By contrast, in conventional education the effectiveness of the teacher as a content expert directly affects the outcomes of the learning process: Poor teaching leads to poor outcomes. In TBL however, the need for explicitly providing content expertise may be significantly reduced by a number of factors. These include high-quality learning materials, challenging applications and effective facilitation. Poor performance of a content expert may be compensated for by high-quality learning materials, a characteristic it has again in common with problem-based learning.[31] Schmidt[32] found that poor tutors in problem-based learning had no negative effect on student achievement if the learning materials, including the problems, were of high quality.The study presented here has several limitations, the most important being that the students' model outlined here represents students' "subjective experiences" with this instructional approach. However, the rather high inter-rater agreements among students who experienced the same course, suggests that their observations are reliable. In addition, students have repeatedly been shown to be rather impartial and reliable observers of their own education.[13]

The second limitation is the reliance on questionnaire data only. It would strengthen the interpretation of the model if some of the variables could have been measured in more unobtrusive ways. It is certainly true that schools collect relevant performance data, such as iRAT scores, tRAT scores, and end-of course achievement scores. However, these data are questionable as measures of the quality of learning going on in a particular course. Do low iRAT scores indicate poor learning or rather a poor match with the learning materials studied? Or do they signify that the teacher misconstrued the ability level of his or her students when he or she developed the test? A more effective, if extremely expensive, approach would be to have independent observers rate the various elements of TBL. Provided their inter-rater reliability would be sufficiently high, such observations may provide an independent window on what happens to the learner in TBL. However, to allow for the kind of analyses conducted for this article, the number of TBL-sessions observed should be at least in the order of 100.[11]

Some readers may consider the use of single items as measures of the constructs in the model a third shortcoming. However, single items as indicators are taken more seriously in social or clinical sciences research than they were a few years ago, in particular if they can be shown to be reliable, such as is the case with our data.[33] In other domains, the use of single items is more common. Indeed, in how many different ways can you ask how much cognitive load a participant experiences in an experiment,[34] how situationally interested a student is in a particular topic,[35] or how severe the pain is which a patient feels?[36]

Where to go from here? Understanding TBL in depth would require two types of studies. One is the experimental approach. Here the complexity of the learning environment would be reduced, so that the various steps can be run through within a reasonable time frame such as 1–2 h. This has been successfully attempted in other complex learning environments.[37] Preparation for TBL would take place under supervision, so that what students learn can be closely monitored. In addition, subparts of the whole would be studied separately. A thought experiment: The tRAT, when students discuss their individual answers to the iRAT, is supposed to increase student's conceptual understanding of what is learned. A design that could test such hypothesis is comparing two groups in terms of knowledge acquired, one of which stops after iRAT, whereas the other continues with tRAT.

The second type of study would be anthropological. Here learning and its context would be studied by participatory observation. Are the assumptions met of those who have developed the approach? Do the processes assumed actually take place? Etc. In addition, products of learning would be studied. For instance, a content analysis of group discussions would reveal to what extent these discussions help in gaining a deeper understanding of the learning.

Acknowledgements

We would like to thank the students for providing us with their feedback.

Declarations

Ethics approval and consent to participate

The IRB of Nanyang Technological University approved this study. Participation was voluntary and anonymous, only group data were analysed.

Consent to publish

N/A (data were collected via routine surveys).

Availability of data and materials

Available upon request from the corresponding author.

Competing interests

No competing interests.

Funding

No funding received.

Authors' contributions

All authors contributed to the conceptual setup of the study.

JR: Wrote the first draft, conducted the data analysis.

PR: Conducted the analysis and provided inputs for the Results section.

MF: Verified the results and provided inputs for the Introduction and Discussion section.

NL: Provided inputs for the Introduction section and verified the Results section.

References

1. Michaelsen LK, Knight AB, Fink LD. *Team-Based Learning: a Transformative Use of Small Groups in College Teaching*. Stylus Pub.; 2004.

2. Haidet, P, et al. Perspective: guidelines for reporting team-based learning activities in the medical and health sciences education literature. *Acad Med* 2012;**87**(3):292–299.
3. Michaelsen LK, Sweet M. The essential elements of team-based learning. *New Dir Teach Learn* 2008;116:7–27.
4. Abdelkhalek, N, et al. Using team-based learning to prepare medical students for future problem-based learning. *Med Teach* 2010;**32**(2):123–129.
5. Parmelee DX, DeStephen D, Borges NJ. Medical students' attitudes about team-based learning in a pre-clinical curriculum. *Med Educ online* 2009:14.
6. Fatmi, M, et al. The effectiveness of team-based learning on learning outcomes in health professions education: BEME Guide No. 30. *Med Teach* 2013;**35**(12):e1608–e1624.
7. Haidet P, Kubitz K, McCormack WT. Analysis of the team-based learning literature: tbl comes of age. *J Excell Coll Teach* 2014;**25**(3–4):303–333.
8. Sisk RJ. Team-based learning: systematic research review. *J Nurs Educ* 2011;**50**(12):665–669.
9. Koles, PG, et al. The impact of team-based learning on medical students' academic performance. *Acad Med* 2010;**85**(11): 1739–1745.
10. McInerney MJ, Fink LD. Team-based learning enhances long-term retention and critical thinking in an undergraduate microbial physiology course. *J Microbiol Biol Educ* 2003;**4**(1):3–12.
11. Nieder, GL, et al. Team-based learning in a medical gross anatomy and embryology course. *Clin Anat* 2005;**18**(1):56–63.
12. Thompson, BM, et al. Team cohesiveness, team size and team performance in team-based learning teams. *Med Educ* 2015;**49**(4):379–385.
13. Marsh HW. Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility. *J Educ Psychol* 1984;**76**(5):707–754.
14. De Grave WS, Dolmans DH, van der Vleuten CP. Tutor intervention profile: reliability and validity. *Med Educ* 1998;**32**(3):262–268.
15. Gijselaers WH, Schmidt HG. Development and evaluation of a causal model of problem-based learning. In: Nooman ZH, Schmidt HG, Ezzat ES, editors. *Innovation in Medical Education: An Evaluation of Its Present Status*. New York: Springer Publishing Co.; 1990. p. 95–113.
16. Van Berkel H, Schmidt HG. On the additional value of lectures in a problem-based curriculum. *Educ Health* 2005;**18**(1):45–61.
17. Hrynchak P, Batty H. The educational theory basis of team-based learning. *Med Teach* 2012;**34**(10):796–801.
18. Shapiro AM. How including prior knowledge as a subject variable may change outcomes of learning research. *Am Educ Res J* 2004;**41**(1):159–189.
19. Bleske, BE, et al. Team-based learning to improve learning outcomes in a therapeutics course sequence. *Am J Pharm Educ* 2014;**78**(1):5.
20. Huitt TW, Killins A, Brooks WS. Team-based learning in the gross anatomy laboratory improves academic performance and students' attitudes toward teamwork. *Anat Sci Educ* 2015;**8**(2): 95–103.
21. Michaelsen LK, Knight AB, Fink LD, editors. *Team-Based Learning: A Transformative Use of Small Groups*. Portsmouth, NH: Greenwood publishing group; 2002.
22. Parmelee, DX, et al. Team-based learning: a practical guide: AMEE Guide No. 65. *Med Teach* 2012;**34**(5):E275–E287.
23. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 1998;**7**(3):301–317.
24. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;**1**(1):30.
25. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model: Multidiscip J* 1999;**6**(1):1–55.
26. Arbuckle JL. *IBM SPSS Amos 23 User's Guide*. Crawfordville, FL: Amos Development Corporation Corporation; 2014.
27. Parmelee DX, Michaelsen LK. Team-based learning: it's here and it WORKS!. *Acad Med* 2010;**85**(11). [1658-1658].
28. Bloom BS. *Human Characteristics and School Learning*. New York: McGraw-Hill; 1976.
29. Near JA, Martin BJ. Expanding course goals beyond disciplinary boundaries: physiology education in an undergraduate course on psychoactive drugs. *Adv Physiol Educ* 2007;**31**(2):161–166.
30. Saye JW, Brush T. Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Etr&D-Educ Technol Res Dev* 2002;**50**(3):77–96.
31. Dolmans, D, et al. Should we choose between problem-based learning and team-based learning? No, combine the best of both worlds!. *Med Teach* 2015;**37**(4):354–359.
32. Schmidt HG. Resolving inconsistencies in tutor expertise research - does lack of structure cause students to seek tutor guidance. *Acad Med* 1994;**69**(8):656–662.
33. Butt, Z, et al. Use of a single-item screening tool to detect clinically significant fatigue, pain, distress, and anorexia in ambulatory cancer practice. *J Pain Symptom Manag* 2008;**35**(1): 20–30.
34. Paas, F, et al. Cognitive load measurement as a means to advance cognitive load theory. *Educ Psychol* 2003;**38**(1):63–71.
35. Ainley M, Hidi S, Berndorff D. Interest, learning, and the psychological processes that mediate their relationship. *J Educ Psychol* 2002;**94**(3):545–561.
36. De Boer, AGEM, et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life?. *Qual Life Res* 2004;**13**(2):311–320.
37. Loyens, SMM, et al. Problem-based learning as a facilitator of conceptual change. *Learn Instr* 2015;**38**:34–42.