ORIGINAL RESEARCH

Inserting an Erroneous Element in the Answer Key of Grouped Team Readiness Assurance Test in Team-Based Learning: A Randomized Controlled Trial in Clinical Toxicology

Asmaa F Sharif [1], Zeinab A Kasemy [2], Nada A Alquraishi[3], Kawther N Alshaikh[3], Anwar H Alfaraj[3], Eman A Alibrahim[3], Zainb M Alfarej[3], Hawra M Alawami[3], Fatma Alzahraa A Elkhamisy [4,5]

[1]Forensic Medicine and Clinical Toxicology Department, Faculty of Medicine, Tanta University, Tanta, Egypt; [2]Public Health and Community Medicine Department, Faculty of Medicine, Menoufia University, Shiben ElKoom, Egypt; [3]College of Medicine, Dar Al-Uloom University, Riyadh, Kingdom of Saudi Arabia; [4]Basic Medical Sciences Department, Faculty of Medicine, King Salman International University, South Sinai, Egypt; [5]Pathology Department, Faculty of Medicine, Helwan University, Cairo, Egypt

Correspondence: Asmaa F Sharif, Forensic Medicine and Clinical Toxicology Department, Faculty of Medicine, Tanta University, Tanta, 31527, Egypt, Tel +966 581639317, Email asmaa.sherif@med.tanta.edu.eg

Purpose: Team-based learning (TBL) is an interactive instructional strategy designed to enhance student engagement. Few studies reported that TBL needs to be more interactive, concerns were raised regarding the insufficient instructor feedback, and students feel that TBL gets less attractive with time. One method of boosting students' interaction and performance is adding an element of challenge or positive stress. Therefore, we hypothesized that inserting an erroneous answer in the answer key would generate an observable improvement in the selected outcomes in terms of better satisfaction, interaction, interest in learning, better academic performance, and better development of competencies compared to traditional TBL.

Methods: This randomized controlled trial aimed to determine if inserting an erroneous element in the answer key of a grouped team readiness assurance test (GRAT) would update TBL and whether this intervention would enhance students' performance and satisfaction.

Results: A total of 55 medical students were recruited (88.7% response rate). Erroneous elements were inserted in the answer key of the experimental group and students enrolled in traditional TBL were considered as controls. Compared to the control group (p < 0.001), the experimental group revealed significantly higher academic performances in GRAT and team evaluation test (TET). Analysis of students' perception of the implemented TBL revealed better perception among the experimental group (33.7 ± 6.4) than the control group (30.1 ± 7.0). Moreover, significantly higher team dynamics were reported among the experimental group than the control group (33.0 ± 6.3 and 27.8 ± 7.6 for both groups, respectively; p = 0.005). The reported advantages were in-depth understanding, easier information retrieval, and development of problem-solving skills. Students considered time and effort as their main limitations.

Conclusion: Adding a few erroneous answers in the GRAT is well perceived by students, enhances their learning competencies and overcomes some TBL challenges.

Keywords: learning environment, motivation, instructional design, student's interaction, academic performance, team-based learning

Introduction

Team-Based Learning (TBL) is a learner-centered interactive instructional strategy designed to enhance students' engagement and interaction.[1] It was originally invented and introduced in higher education in the late 1970s. It is grounded on the constructivist learning theory in which students develop new personal mental frameworks built upon previous knowledge.[2] TBL promotes the development of students' competencies, including self-directed learning, critical thinking, analysis, and application, besides team collaboration and interpersonal skills.[3]

865

TBL is an instructional design that is planned to be carried out over multiple stages. TBL starts with pre-class students' preparation, implements individual and grouped team readiness assurance tests, and ends up with the evaluation and application of the course concepts.[1,4]

The Individual Readiness Assurance Test (IRAT) is the first step in which students are requested to solve a prepared list of Multiple-Choice Questions (MCQs) individually. After that, students are distributed into small teams and solve the same set of MCQs through group discussion until they reach a consensus conclusion and this represents the Grouped Team Readiness Assurance Test (GRAT).[5]

Although TBL was first introduced in business courses, its implementation extended to various educational settings, including medical education.[6] TBL implementation in basic, clinical, and integrated undergraduate medical curricula has increased exponentially during the last few decades.[7,8] Many studies reported the positive impact of TBL in medical education.[1,4] The reported effects varied among studies, from being better to being as effective as the traditional instructional designs to achieving the desired learning outcome.[1]

On the other side, Yang et al reported that traditional TBL, compared to lecture-based learning, is not efficient in attaining the desired learning outcomes in students with poor self-directed learning abilities, and in the acquisition of deep and comprehensive knowledge, with no significant higher achievement of students.[9] Likewise, few studies reported that TBL needs to be more interactive, concerns were raised regarding the insufficient instructor feedback, and students feel that TBL gets less attractive with time.[4,10] Moreover, standardization of TBL implementation and the requirements to repeat it frequently for achieving better outcomes can be challenging. Accordingly, and though TBL is proven to have genuine advantages in medical education, more studies are needed to find innovative ways to potentiate its positive impacts and decrease the few reported challenges.

Most previous studies in the domain of medical education aimed to evaluate TBL in terms of students' perception, achievement, and performances, and very few studies elaborated on the development of the traditional conceptual framework of TBL to boost students' achievement and competencies and overcome the reported challenges. However, these studies had reported positive outcomes compared to traditional TBL.[11,12]

The reported TBL advantages emerge from the challenging nature of this instructional design. One method of boosting students' interaction and performance is adding an element of challenge or positive stress. Rudland et al reported positive correlations between perceived learning and challenges within certain circumstances.[13] A challenge is defined as "Something that needs skills, energy and or determination to deal with or achieve".[14] Positive challenges are eustress, a form of positive emotions that result in better cognitive tasks performance like hypothesis testing.[15] Eustress was reported to affect the motivation,[16] and feeling of enjoying the work.[17,18] Among the reported effects were the enhancement of cognitive memory functions,[19] more motivation, stimulation of excitement, focus enhancement,[20] and promoted performance.[21] Furthermore, eustress increases students' engagement and sensation of self-efficacy.[22]

In the current study, the modified TBL was conducted by changing the correct answer keys in a few questions before requesting the students to collaborate as teams to solve the GRAT. Students work in teams to solve every question and conclude the best answer. Then, they score their answers based on the keys issued by the tutor. Apart from answering the questions by choosing a single best answer, students should seek the erroneous keys and make an argument to defend their choices and why they specifically thought these keys were erroneous. We expect that modifying the TBL, implying the mentioned modification, will generate an observable improvement in students' collaboration, engagement, and interaction which will be reflected as an improvement in their performance at the Team Evaluation Test (TET).

Traditional TBL is grounded in the constructivist theory of education, which is derived from constructivist epistemology. This theory states that "knowledge is an autonomous and subjective construction". The constructivist theory of education is the reinforced form of cognitive learning theory. The learner establishes a long-term memory by integrating the previous experiences and the new information acquired from the surrounding environment. The TBL fits into this theory by adopting its primary four criteria: using a learner-centered approach, enhancing critical thinking and problem-solving skills, and including one tutor who manages many students who actively apply the acquired knowledge through working in teams.[2]

In the current study, the idea of modifying the TBL, besides grounding on the constructivist theory, is built on the stimulus-response theory. The stimulus-response paradigm states that reward enhances optimum response to a specific

stimulus in the learning context. This theory outlined three major elements to enhance the learning: frequency of the used stimuli, the delay time between the response and reward, and the continued linkage between the stimulus and the reward.[23] Modifying TBL by adding more challenging erroneous keys and requesting students to identify these keys immediately and defend their choices redirect the students learning mode from dualism, "ideas are either false or true, and tutors are always right", to the multiplicity in which the true and false may change with the context, and the tutors are not an absolutely true.[24] Elkhamisy and Zidan described adding few erroneous answer-keys to TBL, in the pathology course, as well perceived adjustment based on student's narrative responses.[11]

Therefore, we hypothesized that inserting an erroneous answer in the answer key would generate an observable improvement in the selected outcomes in terms of better satisfaction, interaction, interest in learning, better academic performance, and better development of competencies compared to traditional TBL. The current study aimed to answer the following questions:

- Does inserting erroneous elements in the GRAT answer key become a good reasonable update to the traditional TBL conceptual framework?
- Does inserting erroneous elements in the GRAT answer key enhance students' performances?
- Does inserting erroneous elements in the GRAT answer key satisfy and is well perceived by students?

Material and Methods
Study Design, Setting, and Sampling
A randomized controlled trial was conducted on fourth-year medical students (Phase III of the six-year Bachelor of Medicine and Surgery MBBS Program) from the College of Medicine, Dar Al-Uloom University, between October and November, 2021. TBL was applied as an instructional strategy during the clinical toxicology course taught during the first semester. The clinical toxicology course (two credit hours) is designed to be carried out over 32 contact hours. TBL accounts for 25% weight of formative assessment in the course (eight TBLs). Out of them, four TBLs were re-designed to modified TBL sessions.

Using a computerized random digital method, students who participated in the study were randomly distributed into two groups: control group (traditional TBL: TBL without erroneous answer key) and experimental group (modified TBL: TBL with erroneous answer key). Every student had an equal chance to participate in either group. Blinding was applied from the start of TBL until the end. Baseline comparison of the academic performance between groups was performed, using IRAT, to determine the extent to which the randomization process produced homogeneous groups in terms of relevant background variables.

Random sampling was used in which all consenting students were considered eligible to participate in either group. Sixty-two students are registered in the fourth year. Out of them, two students withdrawn, and the other five students were absent during the TBL sessions, so that, they were excluded from the current study. A total of 55 students were enrolled. Out of them, 25 students were in the experimental group (five teams) and 30 students (six teams) constituted the control group. TBL sessions for both groups were conducted in a separate class for each group.

Ethical Considerations
The current study was conducted after obtaining ethical approval from the Institutional Review Board, College of Medicine, Dar Al-Uloom University (approval number Pro21110001). Responses of all students were handled anonymously according to the principles of the Declaration of Helsinki, 1964, and its later amendments.[25] Participants were briefed regarding the objectives and phases of the study, and each of them provided their written informed consent prior to participating in the study. The informed consent included an agreeing statement for publication of anonymized quotes.

Piloting
Before commencing data collection, a pilot study enrolled 20 participants from fourth-year medical students was carried out. The pilot study aimed to ensure 1) Is sufficient time available in the course for students to work both in class and out

of class to evolve into a cohesive team? 2) Is the teaching faculty prepared to deviate from the traditional TBL-based strategy? 3) Is there administrative "higher ups" support to champion the adopted strategy? 4) Are appropriate session objectives, reading assignments and "content specific learning topics" were created? 5) Are there any potential obstacles or problems that might arise during the work? 6) Is the adopted survey adequate and well prepared?

Feedback from the pilot study showed that students and faculties are well prepared for the modified TBL. A reading guide consisting of different questions sent to the students with the learning material will facilitate students' focus on the essential concepts. Moreover, the time allocated for GRAT and searching the erroneous key was reduced from 30 minutes to 20. Students need ten minutes to complete the survey. Pilot results were excluded from the final results based on the conducted modifications.

Intervention (The modified TBL Method)

The current study was conducted according to the TBL guidelines established by team-based learning collaboration as shown in Figure 1.[26] One week before the planned sessions, the students and faculties were invited to a workshop. During the workshop, they were briefed on the TBL technique and its rationales. Videos were displayed during the workshop to clarify any ambiguities and misconceptions. The TBL conceptual framework was carried out in the following steps:[27]

Team Building

Students were randomly classified into 11 teams of 5 students. Though literature supported that the instructor should select the teams to ensure uniform distribution of students,[28] neither students nor instructors were involved in team selection to avoid selection bias. Teams were permanent to allow better team cohesion. Students were required to introduce themselves to their colleagues so that they could familiarize with other team members. This process helped create diverse teams and ensure proper team dynamics.
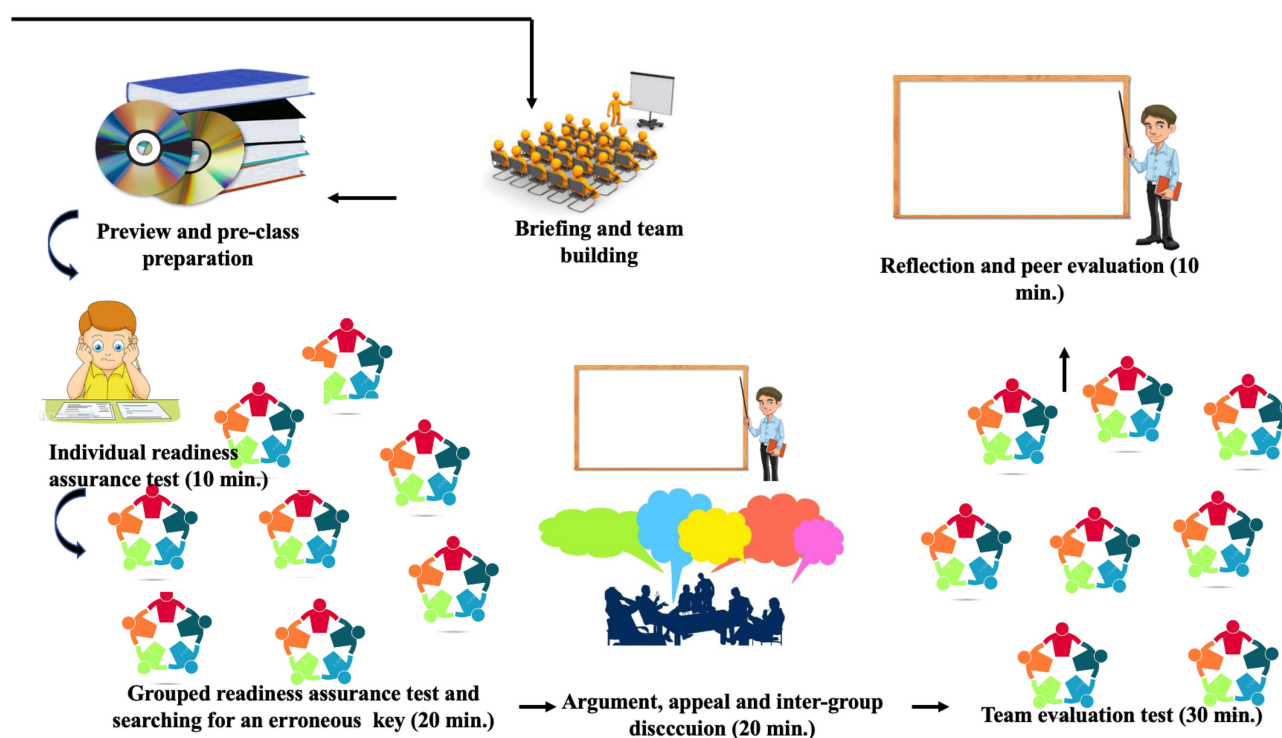
Figure 1 Conceptual framework of modified team-based learning (TBL) method.

Preview and Pre-Class Preparation Stage

Learning material and foundational knowledge were sent to students along with self-directed learning materials one week before the session. Teaching materials were adopted from textbooks that were approved as learning resources in the course specification, including Poisoning and Drug Overdose by Olson et al, and Lippincott's Manual of Toxicology.[29,30] Reading guide, which is a series of questions the students have to answer, was sent for the students to help them identifying the important concepts in the learning material. At this stage, the MCQs used for readiness assurance process were designed by a content expert, then re-evaluated by two faculties experienced in assessment tools, of the same discipline specialty, and proper adjustments were done according to their feedbacks. Questions were typically targeting different levels of Bloom's Taxonomy.

Individual Readiness Assurance Test

IRAT was conducted for the first 10 minutes where each participant student was requested to access the exam link (through Google forms) and solve a list of 11 MCQs, with 4 distractors for each question, individually. Students could not see their grades after submitting their responses.

Grouped Team Readiness Assurance Test

The GRAT was conducted for 20 minutes, and every team was requested to solve the exact same list of MCQs used in IRAT. Group discussions were essential for reaching a consensus on answers to the MCQs. Model key-answer bubble sheets were distributed to all teams where all distractors were covered with bubbles and students of both groups were requested to scratch off the bubbles to check whether their answers were correct or not. In case they find the proper answer different from their selected answer, they should discuss the question again until agreeing on a new answer. Then, they should uncover the bubble next to their newly selected answer to check if this is the true answer or not. This process continues for each question until the team finds that their answer is the true one under the bubble. Teams have to attempt to solve all questions.

Searching for Erroneous Key Answers

Of 11 key-answer sheets, 5 sheets included 1 erroneously answered question in each, thereby maintaining the same set of distractors but altering the key. The five teams of the experimental group were additionally required to identify the question with the erroneous key, select the correct answer for this question from the same list of distractors, and prepare their arguments[11] (Figure 2).

By the end of this step, all teams of both groups got immediate feedback through reaching the correct answers to all questions. Additionally, teams of the experimental group nominated the question with the erroneous key. Bubble sheets were collected for grading. The question with the erroneous key was excluded for both groups and no extra points were counted for experimental group teams to identify the erroneous answer in the GRAT key and explain their rationale to avoid bias in comparing groups.

At the time the students in the experimental group were searching for the erroneous key and preparing their arguments, students in the control group started preparing their arguments and appeals in choosing incorrect answers, and if that was due to vagueness of the question or in readings. All teams, in both groups, were requested to write their appeals and to provide evidence for those appeals from the reading material.

Intergroup Discussion and Appeal

The teams were united and started to present their works. Students' arguments about the questions and/or the answers were presented at this stage. Students of both groups were allowed to present any arguments they wished to. Students of the experimental group presented arguments related to the question with the erroneous key also. The instructor focused on problematic concepts and discussed them thoroughly. Twenty minutes were assigned for this stage.
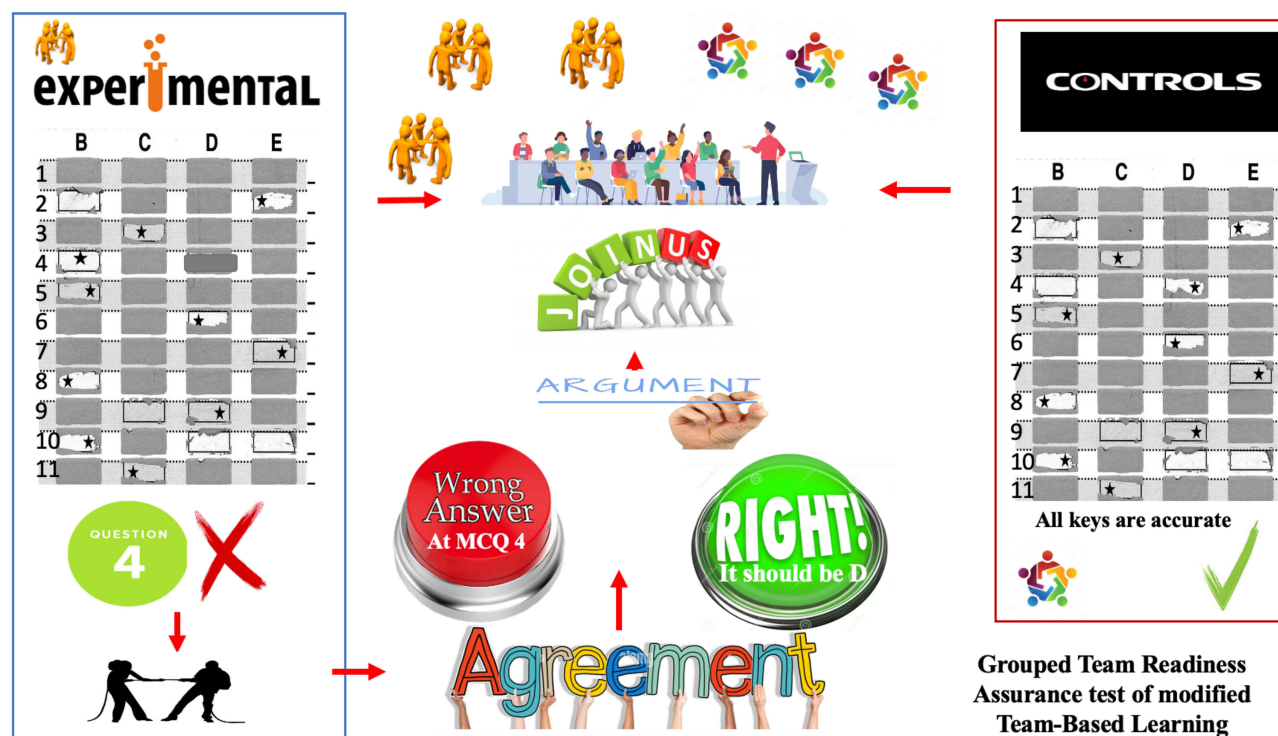
Figure 2 Grouped team readiness assurance test (GRAT) of modified team-based learning (TBL). In the step of GRAT of modified TBL, the students in the experimental group did two tasks: solving the MCQs set as a team and going through a discussion that should end up with extracting one question with an erroneous key. To do that, students should defend their choice and convince each other with that question. When reaching an agreement, students should propose the ideal answer and prepare a written agreement justifying their choice. Then teams are joined to answer all questions and represent their arguments. Taking a look at the bubble sheet depicts that the students in all teams solve the same MCQs with the same set of four distractors. Every team in the experimental group had only one question containing the wrong key. In this example, question number four was wrongly answered with option B. Compared with the control answer key, the appropriate answer is D. Inspecting the answer sheet of both teams alleviates that student answered some questions from the first trials while they could answer some other questions after two or three trials. The question number 4 was excluded from grading for both groups and no extra points were counted for it.

Teams' Evaluation Test

This session started with the distribution of a scenario-based problem with a high difficulty level for every team. Students were previously informed about the nature of the task and the time allocated for it. Team members were provided 30 minutes to analyze and engage in intragroup discussions to solve the problem; responses of the teams were collected for grading after the class. During this stage, students' intragroup discussions were followed by intergroup discussions that were facilitated by the instructor.

Reflection and Peer Evaluation

After the TET session, every student was requested to evaluate the other team members' performance relative to themselves. Anonymized peer evaluations were collected by the instructor, who send each student his/her rating and any comments. Ten minutes were allocated for the feedback where the instructor summarizes the learning objectives and evaluates teams' performances. Students were provided with constructive feedback based on their performance as teams. Furthermore, the instructor provided recommendations regarding the team dynamics and performances.

Data Collection Instruments
Grading
In four TBL sessions (IRAT + GRAT + TET), students were scored considering the average of three tests, each of which was blindly graded out of 10.

Survey (Quantitative Responses)

After the reflection and peer evaluation stage, students were given ten minutes to complete an online survey; the link was sent on Google forms. A structured self-administered survey consisted of three sections was adopted. Each section was scored individually as follows:

(a) The first section comprises five items addressing the students' satisfaction with TBL and average difficulty of TBL. The first subsection evaluates students' satisfaction with TBL in its current form and the traditional lectures. On a five-point Likert scale, the responses ranged from extremely satisfactory, satisfactory, neutral, unsatisfactory, to extremely unsatisfactory. In the second subsection, students were requested to evaluate TBL teaching materials and assessment tools' difficulty level. The pre-session learning materials—IRAT, GRAT, and TET—were graded by students as extremely difficult, difficult, medium, easy, or extremely easy. Then the total difficulty was calculated as the average of all scorings (out of five). Cronbach's alpha calculated for the first section was 0.80.

(b) The second section evaluates students' perception of the TBL in its current form (with and without inserting an erroneous element in the answer key of GRAT). This section consists of nine items with each item's score (1–5) where five is the best and 1 is the least score. Only item number 7 "TBL, in its current form, increased my workload" was scored negatively. The total TBL perception was counted as the sum of each item score where 45 is the best score and 9 is the least score. Questions examining students' perceptions of TBL include whether they prefer TBL to traditional lectures, whether TBL encourages interesting learning, whether TBL enhances self-directed learning (SDL), whether TBL increases analysis and problem-solving skills, whether TBL promotes communication skills, and whether TBL is better organized and aids understanding compared to traditional lectures. Students were asked whether they prefer having access to an electronic device for information during TBL. Responses were scored on a Likert scale; the responses ranged from strongly agree, agree, neutral, disagree, to strongly disagree. Students were asked if the TBL increased their workload and if they would recommend using TBL, in its current form, in other courses. Classification into good, fair, and poor was based on Bloom's cutoff categories.[31] Cronbach's alpha calculated for the second section equals 0.81.

(c) The third section evaluates TBL in terms of team dynamics. This section comprises eight items with scores ranging from 1 to 5 with 5 being the highest and 1 being the lowest. The total score is the sum of scores of all items with 40 being the highest and 8 being the lowest. The questions included in this section were as follows: if the members were encouraged to share their opinions, if they felt it is better to work within a team than to work individually, if they felt they belonged to the right team, if the team actively provoked different viewpoints before deciding on the final answer, if all team members had significantly participated in group discussions, if different viewpoints were respected by different team members, if all team members were attentive during group discussions, and if students assumed that other students went through the readings prior to the session. Classification into good, fair, and poor was based on Bloom's cutoff categories.[31] Cronbach's alpha for this section was 0.85. The overall reliability of the adopted quantitative survey (including all sections) was estimated, and the total Cronbach's Alpha was 0.82.

Qualitative Responses

In the next step, the online narrative feedback of students enrolled in the modified TBL was collected after analyzing the collected data (one week after) to better understand students' perceptions about the modified TBL. Using online Google forms, students were asked to describe their experience and state the advantages, disadvantages, and recommendations to improve TBL in its modified form.

Power Analysis

A statistical power analysis was performed after sample size estimation, based on data from the current study (N = 55), comparing the experimental group to controls. The Effect Size (ES) for GRAT and TET in this study was 1.31 and 3.98, respectively, considered to be large using Cohen's (1988) criteria, with an alpha = 0.05 and sample size = 25 in experimental and 30 in controls, a post-hoc power analysis was conducted with these effect sizes (G Power 3.1) and it

was approximately (1-β) = 0.997 and 1.0, respectively.[32] Thus, the power analysis for a sample size of 55 was adequate for the main objective of this study.

Data Analysis

Quantitative responses of students were analyzed using IBM® SPSS® Statistics version 28 software (IBM Corp., Armonk, NY, USA). Data were expressed as the mean ± standard deviation for quantitative variables and as frequencies (%) for categorical variables. Linear Numeric Response Format was applied. A key phrase "Very unsatisfactory" was at one end and at the other end was "Very satisfactory" or "Strongly disagree" and "Strongly agree". Anchors were provided only at the extremes and no anchors for the middle points. Then the questions were grouped up; hence aggregates of rating scales could be analyzed as continuous data.[33] All tests were conducted using a two-sided significance level of 0.05. Normality tests were conducted, the independent *t*-test was used for normally distributed quantitative variables, and the Mann–Whitney *U*-test was performed for data with skewed distributions. The effect size (Cohen's d) was calculated. The two-way repeated-measures analysis of variance (ANOVA) was performed for the two groups (time-condition interaction). The two groups as a between-subject factor (group) and the three interval measurements as the within-subject factor (time) were considered.

Qualitative narrative responses of students were analyzed iteratively using an inductive approach and thematic analysis. Responses were analyzed until the saturation point was reached (no new themes emerged).[34] Two investigators conducted three cycles of coding.

Rigor
Reliability
The major themes were continuously extracted until no new themes emerged to confirm the validity of the current study, especially the qualitative data analysis. Responses were extracted and categorized under three major themes: advantages, limitations/disadvantages, and recommendations. For each theme, underlying categories were identified along with their corresponding codes and quotations from students' responses.

Responses were obtained and blindly coded by two independent investigators. A coding book was developed where the agreed definitions and inclusion and exclusion criteria were documented. The interrater error was calculated using the formula "Interrater reliability = number of agreed codes/number of agreed codes + number of disagreed codes".[35] Out of 21 codes, 14 codes were agreed with an overall good reliability of 0.7.

Validity
We adopted the following strategies to confirm the validity of the conducted study:

(i) Students were randomly assigned to groups and teams; scores were calculated blindly, and data analysis was conducted without group nominations.
(ii) Triangulation of investigators (Two investigators [AS] and [FE] conducted the data extraction).
(iii) All student responses were analyzed continuously.
(iv) All components of the learning environment, including learning resources, assessment tools, instructors, classrooms, and timing of the TBL, were shared between the two groups.

Results

Fifty-five out of 62 students participated in this study, which implies that the response rate was 88.7%. Besides two students who withdrew, the other five excluded students were absent in one or more sessions. Among the total study participants, 30 (54.5%) were females and 25 (45.5%) were males. The mean age of study participants was 21 ± 0.6 years, and there was no significant difference in the mean age between males and females.

As shown in Table 1, both groups revealed scores that increased as the TBL stages advanced, with the lowest scores recorded in IRAT; the scores increased in GRAT and reached the maximum in TET. Both groups showed similar means of IRAT scores (4.3 ± 2.2 for the experimental group versus 4.5 ± 2.3 for the control group) with no significant differences between the two mean scores.
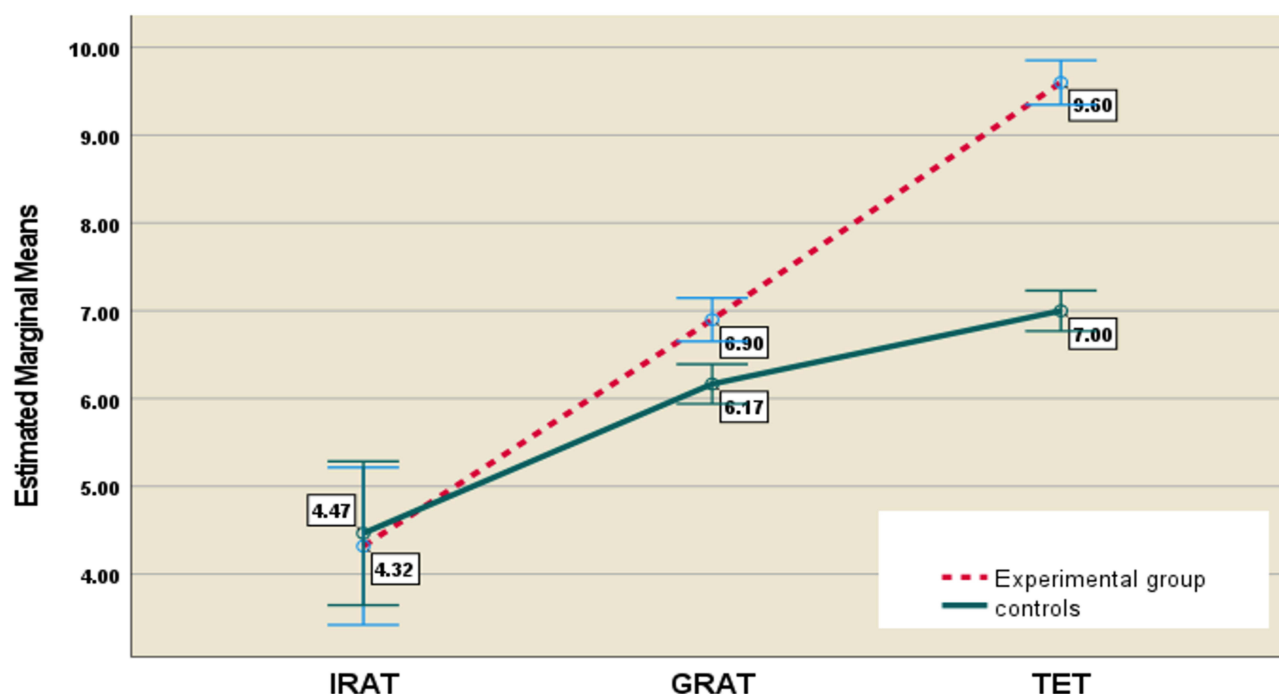
Table 1 Distribution of the Studied Groups Regarding Individual Readiness Assurance Test (IRAT), Grouped Team Readiness Assurance Test (GRAT) and Team Evaluation Test (TET)

Students Performance	Experimental (No.=25) (Mean ±SD)	Controls (No.=30) (Mean ±SD)	Effect Size (95% CI)	Test of Sig	P value
• **Individual Readiness Assurance Test** (out of 10)	4.3±2.2	4.5±2.3	−0.06 [(−0.59)-(0.46)]	0.38	0.701
• **Grouped Team Readiness Assurance Test** (out of 10)	6.9±0.5	6.1±0.7	1.18 [0.60–1.76]	4.39	<0.001*
• **Team Evaluation Test** (Out of 10)	9.6±0.6	7.0±0.7	4.13 [3.17–5.07]	15.25	<0.001*
• **Total average evaluation of the three tests** (Out of 10)	6.9±0.8	5.9±1.0	1.13 [(0.55)-(1.70)]	4.18	<0.001*

Note: *: significant.

Alternatively, the experimental group demonstrated considerably higher scores in GRAT and TET exams, in addition to the three tests' combined average (p < 0.001). The difference in scores was particularly obvious in the TET where the mean score of the study group was 9.6 ± 0.6 compared to 7.0 ± 0.7 in the control group. Moreover, the two-way ANOVA revealed that the difference between the two groups was statistically significant, as indicated by the effect of the experiment, using the between-subject factor [F(1.08, 52.27) = 133.08, p < 0.001, η2 = 0.715)]. The interaction between tests and groups seemed to be linear [F(1, 53) = 21.76, p < 0.001, η2 = 0.769]. The main effect of the experiment on scores across tests was statistically significant [F (1, 53) = 17.49, p < 0.001, η2 = 0.248] (Figure 3).

As shown in Table 2, both study groups considered learning materials, RAT, and TET at similar difficulty levels (from medium to difficult) with no significant difference observed between the groups (p > 0.05). Similarly, both groups rated the TBL in its current form (modified and original) as more satisfactory than traditional lectures. Students in the experimental group revealed higher levels of satisfaction with TBL (mean score = 4.2 ± 1.2) compared to those in the control group (mean score = 4.0 ± 1.0). However, this difference was not statistically significant (p = 0.581).

Figure 3 Individual Readiness Assurance Test, Grouped Team Readiness Assurance Test and Team Evaluation Test over the studied participants.

Table 2 Distribution of the Studied Groups Regarding Satisfaction and Difficulty of Team-Based Learning (Original and Modified TBL)

Student's Satisfaction with TBL and TBL Difficulty Level	Experimental (No.=25) (Mean ±SD)	Controls (No.=30) (Mean ±SD)	Effect Size (95% CI)	Test of Sig	P value
● Rate traditional lectures according to your satisfaction	3.6±1.2	3.6±1.1	–	–	–
● Rate TBL, in its current form, according to your satisfaction	4.2±1.2	4.0±1.0	−0.15 [(−0.38) -(0.68)]	0.55	0.581
● How you found the difficulty of the learning material	2.8±0.8	3.0±0.9	−0.26 [(−0.79-(0.27)]	0.96	0.338
● How you found the difficulty of the individual test?	3.3±0.9	3.2±0.9	0.09 [(−0.43) -(0.62)]	0.34	0.729
● How you found the difficulty of the case at the end of the session?	3.0±0.8	3.2±1.09	−0.12 [(−0.65) -(0.40)]	0.46	0.646
● Total difficulty	3.05±0.7	3.1±0.9	−0.11 [(−0.64) -(0.41)]	0.43	0.668

Table 3 highlights students' perception of TBL in both groups. Participants in the experimental group revealed an overall better perception of TBL than those in the control group (mean score = 33.7 ± 6.4 versus 30.1 ± 7.0, respectively) (p = 0.060). Approximately half of the students (48%) who underwent a modified TBL exhibited a good perception compared to 26.7% of students who enrolled in the original TBL. Students considered the modified TBL to be more interest-stimulating (4.3 ± 0.7) than the original TBL (3.7 ± 1.1; p = 0.023). Moreover, students considered modified TBL as being capable of enhancing their abilities to SDL (4.1 ± 1.0) more than the original TBL (3.5 ± 1.1; p = 0.033). Students in the experimental group perceived the modified TBL as more supportive toward enhancing their communication skills (4.2 ± 1.1) as compared to their colleagues in the control group (3.6 ± 1.2; p = 0.048). Students exposed to

Table 3 Distribution of the Studied Groups Regarding Perception of Team-Based Learning (Original and Modified TBL)

Students' Perception of TBL	Experimental (No.=25) (Mean ±SD)	Controls (No.=30) (Mean ±SD)	Effect Size (95% CI)		Test of Sig	P value
● I prefer TBL, in its current form, compared to the traditional theoretical lecture	3.9±1.2	3.5±1.0	0.33 [(−0.19-(0.87)]		1.24	0.217
● TBL, in its current form, stimulates interest in learning	4.3±0.7	3.7±1.1	0.63 [0.08–1.17]		2.34	0.023*
● TBL, in its current form, enhances my ability to self-directed learning when I studied the learning material before the session	4.1±1.0	3.5±1.1	0.59 [0.04–1.13]		2.18	0.033*
● TBL, in its current form, increased my ability to analyze and solve skills	4.2±0.9	3.8±1.2	0.41 [(−0.12)-(0.95)]		1.53	0.130
● TBL, in its current form, promotes my communication skills	4.2±1.1	3.6±1.2	0.54 [0.004–1.08]		2.02	0.048*
● I feel learning using TBL, in its current form, is more organized and helps easier understanding than traditional lecture	3.9±1.3	3.3±1.2	0.46 [(−0.07) -(1.0)]		1.71	0.093
● TBL, in its current form, increased my workload	2.4±1.2	2.3±1.1	0.05 [(−0.47)-0.58]		0.20	0.839
● I recommend implementing TBL, in its current form, in other courses	2.8±0.8	2.9±0.9	−0.06 [(−0.59)-(0.46)]		0.23	0.813
● I would have liked access to an electronic device for information.	3.7±1.0	3.4±1.1	0.27 [−0.26) -(0.80)]		1.0	0.321
Total TBL Perception (out of 45)	33.7±6.4	30.1±7.0	0.53 [(−0.01) -(1.08)]		1.92	0.060
● Good (no, %)	12　48.0	8　26.7	-		4.33	0.115
● Fair	10　40.0	12　40.0				
● Poor	3　12.0	10　33.3				

Note: *: significant.

modified TBL had a better perception of other items than those enrolled in original TBL, such as improving students' analytical and problem-solving skills and aiding in easier understanding, although these differences did not reach statistical significance. Students in both groups considered TBL to increase their workloads in a similar manner and recommended the implementation of TBL in other courses.

Team dynamics are displayed in Table 4. Students in the experimental group revealed significantly higher total team dynamics scores (33.0 ± 6.3) than those in the control group (27.8 ± 7.6; p = 0.005). Most students who underwent modified TBL (68%) demonstrated good team dynamics compared to 36.7% of students who enrolled in original TBL. All items of the survey were scored higher in the experimental group than in the control group. However, significantly better dynamics were reported in the experimental group than in the control group for a few items. These items included students' belief in performing well within a team than working individually (3.9 ± 1.1 for the experimental group versus 3.3 ± 1.1 for the control group, p = 0.017), the team actively eliciting multiple viewpoints before deciding on a final answer (4.2 ± 0.9 for the experimental group versus 3.5 ± 1.1 for the control group, p = 0.035), and efforts of team members to participate in discussions (4.2 ± 0.8 for the experimental group versus 3.5 ± 1.1 for the control group, p = 0.018). Other items that scored significantly higher in the experimental group than in the control group were as follows: consideration of varied viewpoints of team members (p < 0.001) and consistency in paying attention during group discussions (p = 0.007).

Regarding the narrative students' perceptions about the modified TBL, testing the inter-reliability of the obtained responses among seven randomly chosen responses (28% of qualitative responses) revealed adequate agreements with a good overall reliability score of 0.8 (Table 5).

Descriptions and definitions of students' responses along with some quotes are shown in Table S1. Three major themes have emerged from analyzing narrative students' responses, in terms of advantages and disadvantages of the modified TBL, besides the recommendations to improve the modified TBL. The themes with subsequent categories and codes were listed as follows:

Table 4 Distribution of the Studied Groups Regarding Student's Dynamics During Conduction of Team-Based Learning (TBL)

TBL Team Dynamics	Experimental (No.=25) (Mean ±SD)		Controls (No.=30) (Mean ±SD)		Effect Size (95% CI)	Test of Sig	P value
● Team members encouraged one another to express their opinions	4.2±1.0		3.8±1.0		0.37 [(−0.16) -(0.90)]	1.37	0.174
● I feel it is better to work with a team than working individually	3.9±1.1		3.3±1.1		0.66 [(0.12) -(1.21)]	2.47	0.017*
● I feel that I worked with the right team	4.2±0.9		3.6±1.2		0.47 [(−0.07) -(1.0)]	1.73	0.088
● My team actively elicited multiple points of view before deciding on a final answer.	4.2±0.9		3.5±1.1		0.58 [(0.04) -(1.12)]	2.16	0.035*
● All team members made an effort to participate in discussion.	4.2±0.8		3.5±1.1		0.66 [(0.11) -(1.20)]	2.44	0.018*
● Different points of view were respected by team members	4.3±0.8		3.4±1.1		0.91 [0.35–1.46]	3.41	<0.001*
● All team members consistently paid attention during group discussions	4.2±1.0		3.4±1.1		0.75 [(0.20) -(1.30)]	2.78	0.007*
● Students did read the readings prior to session	3.8±0.9		3.4±1.1		0.34 [−0.19) -(0.88]	1.27	0.206
Total dynamics (out of 40)	33.0±6.3		27.8±7.6		0.73 [0.17) -(1.27)]	2.69	0.005*
● Good (no, %)	17	68.0	11	36.7	-	5.95	0.064
● Fair	6	24.0	11	36.7			
● Poor	2	8.0	8	26.7			

Note: *: significant.

Table 5 Inter-Rater Reliability Among Randomly Chosen Sample of Studied Participants (n = 7)

Participant ID	Investigator I		Investigator II	
	Number of Extracted Codes	**Lines**	**Number of Extracted Codes**	**Lines**
S1	3	1–3, 4, 6	2	1–3,6
S9	3	27–29, 30,31	3	27–29, 31, 32
S11	2	35–36,38	2	38, 39
S17	3	50–52, 53, 55	2	50–51, 55
S20	3	62, 63, 64	2	62, 64
S22	2	66–67, 68	2	66–67, 68
S24	3	71–73, 74, 75–76	3	71–73, 74, 75–77

Notes: The interrater error was calculated using the formula "Interrater reliability = number of agreed codes/number of agreed codes + number of disagreed codes". Out of 21 codes, 14 codes were agreed, 7 codes were not, with an overall good reliability of 0.7.

Theme 1. The Modified TBL Has Diverse Advantages from Students' Perspectives

1. The modified TBL improved students' knowledge and cognition in terms of increased focus on subject details, in-depth understanding, and longer retention with easier information retrieval.
2. The modified TBL enhanced students' intellectual skills, including problem-solving and critical thinking, and allowed better use of higher cognitive functions.
3. The modified TBL allowed students to upgrade their general transferrable skills, including teamwork, time-management, leadership and self-reflection.
4. The modified TBL positively influenced students' values and attitudes. They described a state of interaction with better engagement, fun and interesting, socializing, motivating, more thrilling, challenging and competitive, more confident and self-efficient, not boring, and receiving constructive feedback.

One sound said, "I had to focus on the details of each question and the distractors presented to discover where the erroneous answer is". Another sound mentioned "I can remember the lecture more easily after the discussion with the instructor about potential erroneous answers of the key". Sound 1 said "We analyzed the new case and evaluated the information present in it in a better light of what we learnt deeply during the previous steps", Sound 17 mentioned "It is interesting, and I didn't feel bored easily" Sound 18 said "I can remember the lecture more easily after the discussion with the instructor about potential erroneous answers of the key".

Theme 2. The Modified TBL Encountered a Few Disadvantages/Limitations from Students' Perspectives

1. Some teamwork dynamics were negatively perceived including Unresponsive passive team members and Dominant behavior of some members.
2. Limited time and high efforts were the leading limitations for the modified TBL

Students stated, "Some of the team members were not cooperative, they were mere observers to our discussions", "Time allocated for the discussion was insufficient", and "It needed too much effort."

Theme 3. Students' Recommendations to Improve the Modified TBL

1. Students recommended intrateam and inter-team cooperation and interaction to improve teamwork dynamics.
2. Students recommended the use of electronic clickers and extension of duration for the modified TBL.
3. Students recommended receiving personalized feedback.

4. Students recommended generalization of the modified TBL by allowing for more frequent sessions.

Some Students quotes were: "Some team members need to be more interactive; a strategy should be done to push them to have a role in the team in the MCQ & in the cases", "I recommend exchanging our results with other team results and discuss them before announcing them in front of all teams and the tutor", "I felt we need more time to finish our discussion and select our answers confidently", and "I prefer getting additional feedback for my own responses."

Discussion

The current study examined how students' performance and perception of TBL is impacted by modifying the traditional TBL by adding an erroneous element in the GRAT answer key. It revealed that this modification is well perceived by students and enhances their satisfaction, cognition, skills, attitude, and academic performance.

Few trials using simulated patients aimed to modify the TBL conceptual framework to increase its outcomes.[12] Elkhamisy and Zidan suggested modifying the TBL, as part of the pathology course, by adding erroneous element questions that varied 10 times between teams. Their study demonstrated a better perception of the modified TBL, and students denied any feeling of confusion because of this modification. However, Elkhamisy and Zidan only assessed the student's perception qualitatively in the basic medical sciences context without assessing students' performances or using a control group.[11]

The current study demonstrated higher students' achievement in the GRAT phase compared to the IRAT phase in both the control and experimental groups of students. Students learned from one another during their discussions in the GRAT phase. This result is consistent with the results of a meta-analysis conducted by Ngoc et al who investigated students' performances in the GRAT phase compared to the IRAT phase.[5]

Moreover, according to the present study, students' academic performances are significantly improved by the modified TBL compared to the traditional TBL as proved by higher TET results among participants in the experimental group despite insignificant variations in the IRAT score between study groups. The positive influence of TBL on students' performances compared to traditional lectures is consistent with Parmelee et al who reported equivalent or higher performances due to TBL practicing.[1]

Based on the control-value theory of achievement emotions, several factors might impact students' academic performances, including the positive and negative emotions toward learning.[36,37] Students' enjoyment of the learning process has a positive impact on their academic performances while their finding the learning process boring has a negative effect on their academic performances.[38] The positive emotional effect is mediated through intrinsic and extrinsic motivation. Intrinsic motivation implies building positive emotions in students toward themselves and their learning, which becomes autonomous. This state of autonomy motivation can be fostered by activities that are enjoyable and challenging to students, enhancing their feeling of curiosity and competence (ie, self-efficacy).[39] According to the self-determination theory, intrinsic motivation has a higher impact on students' academic performances.[40]

Although traditional TBL has been reported to have a relative boosting effect on motivation because it contains elements that stimulate positive emotions and motivation in students,[41] other studies have reported that students who consider traditional TBL become less interested after the first two to three sessions and they recommend that TBL should be more interactive.[4,10] In this study, the experimental group perceived modified TBL as significantly more interest-stimulating compared to the traditional TBL. Furthermore, students considered the modified GRAT step to be fun and interesting, socializing, motivating, challenging, and competitive. Students also described their attitudes during inter-group discussions as being more interactive and fun with better engagement. Students felt more confident and self-efficient when they were able to present their arguments against the instructor's answers and defend their choices. They described the search for the erroneous answer key as being more thrilling, and they were more focused and attentive during such sessions. All the above-mentioned positive perceptions that accompany the modified TBL promote higher states of motivation and positive emotions toward the learning process; hence, it is associated with better performances compared to the traditional TBL. Elkhamisy and Zidan reported a similar finding in which most students felt more motivated, challenged, and self-efficient in the modified TBL compared to the conventional TBL they used to practice.[11]

Another factor that impacts students' performances is the extent of their cognitive engagement in learning, which is the psychological state of putting significant effort to understand the topic at hand that persists over a prolonged period.[42] It can be fostered partially by creating the autonomy-enhancing learning environment mentioned previously.[43] The traditional TBL was reported to promote variable fluctuating levels of cognitive engagement among students.[42] In the current study, students who enrolled in the modified TBL in spite of the autonomy-enhancing environment reported interesting GRAT with more focus and attention. They also reported greater knowledge acquisition, deeper under-standing, longer retention with easier information retrieval in the intra-group discussions. A deeper and more detailed review of the questions was done in that step. If we considered the significant positive correlation between cognitive engagement and performance, we would conclude that the modification carried out in TBL, in the form of adding an erroneous key, appears to upgrade students' performances.[42] Similarly, Elkhamisy and Zidan also reported more focus on the details of the subject compared to the traditional TBL.[11]

In the current study, students who enrolled in the modified TBL reported higher abilities to analyze and solve problems than those who enrolled in the traditional TBL. Students of the modified TBL experienced higher development of their problem-solving and critical thinking skills as advantages of the modified GRAT phase; they had to use these skills to evaluate the key given to them to discover the erroneous answer, choose a correct one, and defend it, besides the traditional use of these skills in the TET phase. Many studies reported enhanced students' learning gains and achieve-ment by incorporating tasks with a high level of cognitive challenge stimulating their higher-order cognitive skills.[44] Cognitive challenges require higher-order thinking skills like analysis, evaluation, problem-solving, and critical thinking skills.[45,46] Consistent with the results of the current study, Elkhamisy and Zidan reported that their students used critical thinking skills more with the modified TBL.[11]

The current study highlights how students value the well-designed case-based scenarios used in TBL and to what extent they foster higher cognitive skills. In medical education, the effective instructional tool should be relevant to future practice.[47] Considering the variation in TBL task designs, which is reflected in the variable learning effectiveness, there is a need to integrate the components of higher-order thinking, including an analysis of a situation, reasoning, and reflective critical thinking to come up with a judgment in the TBL design.[48]

Another advantage of the modified TBL that we identified in the current study is the significantly higher enhancement of the students' SDL abilities. Similarly, literature reported that the traditional TBL significantly enhances students' SDL capabilities and students' learning satisfaction compared to traditional lectures.[49] SDL is a process in which learners play a dominant role in their learning process. Learners have control over their learning in terms of initiating it; determining their learning needs, goals, and resources; applying suitable learning strategies; and evaluating their learning outcomes.[50] SDL is crucial to the lifelong learning competency that needs to be perfected by medical graduates.[51] Moreover, SDL in health professions education is associated with a moderate improvement in the knowledge domain and enhanced academic performance compared to traditional teaching methods, which is in accordance with the findings of the current study.[52,53] Medical educators should facilitate the acquisition of the SDL skills by students so they can master the use of the SDL process.[54]

Another critical advantage of the modified TBL identified in the current study is the positive influence of modifying TBL on fostering team dynamics. The current study revealed that the modified TBL significantly boosted team dynamics compared to the traditional TBL. Students reported higher scores in several items, including team member efforts to participate in discussion, expression of multiple viewpoints of team members before arriving at a unanimous decision, and improved teamwork. The modified TBL also significantly promoted communication skills of students compared to the traditional TBL. This might be due to the challenge offered to students, which propelled them to collaborate and engage in discussions to "win over" the instructor. They had to listen to each other and were united by a single common vision. Moreover, students in the modified TBL group reportedly developed leadership skills to a higher degree. In effective teams, shared leadership, in which members work collectively to facilitate group processes and improve performance, is the dominant leadership style.[55]

The importance of teamwork dynamics is crucial, especially in health-care education. It is involved in the core of the professionalism competency.[56] The characteristics of great teams include maintaining high energy, ensuring that each

team member contributes in an equally brief manner, building interpersonal relationships among team members, and improving interactions with other teams.[57]

Getting all team members to engage in teamwork is one of the challenges in any team-working context.[58] Apparently, the modified TBL conducted in the current study overcame this challenge to some extent. However, a few students reported that some of their team members were not interactive and some portrayed dominant behavior. Students recommended more cooperation and interaction between team members. This may be fostered by designing an orientation on the developmental phases of teamwork and apprising students of the benefits of getting involved in teamwork to enhance motivation and engagement.[39,58]

Though the anchored challenging nature of the traditional TBL presented in the different stages of TBL, the proposed TBL modification is assumed to be an extra stimulation of these challenges. Modifying the traditional TBL framework by adding erroneous answers in the GRAT and asking students to discover and argue them presents a learning challenge for students.[11,12]

Despite several advantages of the modified TBL elaborated by the current study, there were a few disadvantages reported by students who enrolled in it. Hence, more efforts are required, and more time should be allocated to students to answer the questions in GRAT. The TET steps were the main reported disadvantages. Nevertheless, most students enrolled in the modified TBL recommended the generalization of this instructional design in the clinical toxicology course. Students' tendency to repeat interesting interactive learning activities is not limited to the TBL but extends to involve other similar instructional strategies.[59,60]

Introducing new concepts or change in traditional teaching strategies is challenging and is faced chiefly by significant resistance.[61] While most of previous studies elaborated on the need for developing traditional teaching strategies, few of them outlined clear procedures and steps to do.[62] The current study endorsed a novel way to modify the TBL conceptual framework, based on theoretical, pedagogical backgrounds. Implementing the proposed modified TBL offers diverse positive implications in medical education. Apart from solving the problems of student's amotivation and boredom while exposed to traditional instructional strategies, it offers an environment for creativity and innovation. Moreover, it allows better working memory management and increases cognitive brain functions. Involving medical students in a sequence of complex learning tasks, individual and grouped tests, and SDL activities, which are available in the proposed modification, is promising and substantially enhances all learning domains.[63]

Strength, Limitations, and Recommendations

The current study endorses a novel instructional design to assess GRAT phase modification of the TBL. It is the first study in literature that quantitatively assesses students' perceptions of the modified TBL and their resultant academic performances in medical education. Furthermore, the current study implements the modified TBL in a clinical medical science context.

The study's main limitations stem from the fact that it was limited to one discipline and a single institution. Future studies in multiple centers and on larger cohort of students should be conducted. Moreover, comparative studies implemented on repeating TBL sessions need to be performed.

Conclusion

Modifying the GRAT phase of the TBL by including a few erroneous answers in the answer keys of GRAT as well as asking students to identify, argue, and correct them is an innovative modification in the TBL conceptual framework. It is well perceived by students and enhances their learning competencies, including students' interaction, enjoyment, focus, knowledge, and skills. Moreover, the modified TBL is associated with improved academic performances, which might overcome some reported TBL challenges and boost the learning outcomes.

Abbreviations

ANOVA, Analysis of Variance; ES, Effect Size; GRAT, Grouped Team Readiness Assurance Test; TBL, Team-Based Learning; IRAT, Individual Readiness Assurance Test; MCQs, Multiple-Choice Questions; SDL, Self-Directed Learning; TBLC, Team-Based Learning Collaboration; TET, Team Evaluation Test.

Institutional Review Board Statement

Data Sharing Statement

The data analyzed in the current study are available upon reasonable request from the corresponding author.

Informed Consent Statement

Written informed consent was obtained from every participant prior to their inclusion in the current study, including a consent to publish anonymized quotes.

Acknowledgments

Author Contributions

All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article was submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors report no competing interests (financial or non-financial) related to this work.

References

1. Parmelee D, Michaelsen LK, Cook S, Hudes PD. Team-based learning: a practical guide: AMEE Guide No. 65. *Med Teach*. 2012;34(5):e275–e287. doi:10.3109/0142159X.2012.651179
2. Hrynchak P, Batty H. The educational theory basis of team-based learning. *Med Teach*. 2012;34(10):796–801. doi:10.3109/0142159X.2012.687120
3. Gopalan C, Fox DJ, Gaebelein CJ. Effect of an individual readiness assurance test on a team readiness assurance test in the team-based learning of physiology. *Am J Physiol Adv Physiol Educ*. 2013;37(1):61–64. doi:10.1152/advan.00095.2012
4. Reimschisel T, Herring AL, Huang J, Minor TJ. A systematic review of the published literature on team-based learning in health professions education. *Med Teach*. 2017;39(12):1227–1237. doi:10.1080/0142159X.2017.1340636
5. Ngoc PN, Cheng CL, Lin YK, Wu MS, Chu JS, Tang KP. A meta-analysis of students' readiness assurance test performance with team-based learning. *BMC Med Educ*. 2020;20(1):1–9. doi:10.1186/s12909-020-02139-9
6. Mennenga HA, Smyer T. A model for easily incorporating team-based learning into nursing education. *Int J Nurs Educ Scholarsh*. 2010;7(1). doi:10.2202/1548-923X.1924
7. Zgheib NK, Simaan JA, Sabra R. Using team-based learning to teach pharmacology to second year medical students improves student performance. *Med Teach*. 2010;32(2):130–135. doi:10.3109/01421590903548521
8. Warrier KS, Schiller JH, Frei NR, Haftel HM, Christner JG. Long-term gain after team-based learning experience in a pediatric clerkship. *Teach Learn Med*. 2013;25(4):300–305. doi:10.1080/10401334.2013.827975
9. Yang L, Jiang L, Xu B, et al. Evaluating team-based, lecture-based, and hybrid learning methods for neurology clerkship in China: a method-comparison study. *BMC Med Educ*. 2014;14:1–7.
10. Doshi NP. Effectiveness of team-based learning methodology in teaching transfusion medicine to medical undergraduates in third semester: a comparative study. *Asian J Transfus Sci*. 2017;11(2):87. doi:10.4103/ajts.AJTS_123_16
11. Elkhamisy FA, Zidan AH. Intentional adding of a wrong element in the answer-key for the team readiness assurance test (Abstract). Virtual Conference Book of Abstracts; 2021: 798. Available from: https://amee.org/getattachment/Conferences/AMEE-2021/Programme/AMEE-2021-Abstract-Book-V2.pdf. Accessed August 3, 2022.
12. Fernandes AK, Ecklar P, Rundell K, Luster G, Cavalcanti M. Integrating simulated patients in TBL: a strategy for success in medical education. *Med Sci Educ*. 2019;29(2):383–387. doi:10.1007/s40670-019-00727-z
13. Rudland JR, Jaye C, Tweed M, Wilkinson TJ. Relationships among perceived learning, challenge and affect in a clinical context. *BMC Med Educ*. 2021;21(1):1–10. doi:10.1186/s12909-021-02574-2
14. Macmillan. *The Macmillan Dictionary of Quotations*. Prebenna D, ed. Macmillan Education; 1989. Available from: https://www.macmillandictionary.com/dictionary/british/challenge_1. Accessed April 18, 2022.

15. Nadler RT, Rabi R, Minda JP. Better mood and better performance: learning rule-described categories is enhanced by positive mood. *Psychol Sci*. 2010;21(12):1770–1776. doi:10.1177/0956797610387441

16. Selye H. Stress without distress. In: *Psychopathology of Human Adaptation*. Springer; 1976:137–146.

17. Hargrove MB, Hargrove MB, Nelson DL, Cooper CL. Generating eustress by challenging employees: helping people savor their work. *Organ Dyn*. 2013;2013(42):61–69. doi:10.1016/j.orgdyn.2012.12.008

18. Simmons BL, Nelson DL. Eustress at work: the relationship between hope and health in hospital nurses. *Health Care Manage Rev*. 2001;26(4):7–18. doi:10.1097/00004010-200110000-00002

19. Cahill L, Gorski L, Le K. Enhanced human memory consolidation with post-learning stress: interaction with the degree of arousal at encoding. *Learn Mem*. 2003;10(4):270–274. doi:10.1101/lm.62403

20. Girzadas JDV, Delis S, Bose S, Hall J, Rzechula K, Kulstad EB. Measures of stress and learning seem to be equally affected among all roles in a simulation scenario. *Simul Healthc*. 2009;4(3):149–154. doi:10.1097/SIH.0b013e3181abe9f2

21. Mills H, Reiss N, Dombeck M. Types of stressors (eustress vs. distress); 2008: 2010.

22. Mesurado B, Cristina Richaud M, José Mateo N. Engagement, flow, self-efficacy, and Eustress of University Students: a cross-national comparison between the Philippines and Argentina. *J Psychol*. 2016;150(3):281–299. doi:10.1080/00223980.2015.1024595

23. Taylor DCM, Hamdy H. Adult learning theories: implications for learning and teaching in medical education: AMEE Guide No. 83. *Med Teach*. 2013;35(11):e1561–e1572. doi:10.3109/0142159X.2013.828153

24. Perry JWG. *Forms of Intellectual and Ethical Development in the College Years: A Scheme*. Jossey-Bass Higher and Adult Education Series. ERIC; 1999.

25. PP Rickham. Human experimentation. Code of ethics of the world medical association. Declaration of Helsinki. *Br Med J*. 1964;2(5402):177.

26. TBLC. Team-Based Learning Collaborative. Getting Started with TBL; 2022. Available from: https://teambasedlearning.site-ym.com/page/started. Accessed May 7, 2022.

27. Sibley J, Spirid S. Introduction to team-based learning; 2014. doi:10.1186/s12909-016-0838-3.

28. Brickell LCJL, Porter LCDB, Reynolds LCMF, Cosgrove CRD. Assigning students to groups for engineering design projects: a comparison of five methods. *J Eng Educ*. 1994;83(3):259–262. doi:10.1002/j.2168-9830.1994.tb01113.x

29. Olson KR, Anderson IB, Benowitz NL, et al. *Poisoning & Drug Overdose*. Lange Medical Books/McGraw-Hill; Vol. 13, 2007.

30. Lynch JJ. *Lippincott's Manual of Toxicology*. Lippincott Williams & Wilkins; 2012.

31. Alzahrani MM, Alghamdi AA, Alghamdi SA, Alotaibi RK. Knowledge and attitude of dentists towards obstructive sleep apnoea. *Int Dent J*. 2021;1–7. doi:10.1016/j.identj.2021.05.004

32. Cohen J. *Statistical power analysis for behvioral sciences*. 2nd Ed. New York: NY: Routledge; 1988. doi: 10.4324/9780203771587

33. Franc J. Three ways to analyze likert scales - Avoiding the likert crush - The latest news in research statistics | STAT59; 2020. Available from: https://www.stat59.com/blog/2020/6/three-ways-analyze-likert-scales/. Accessed April 10, 2022.

34. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*. 2004;24(2):105–112. doi:10.1016/j.nedt.2003.10.001

35. Miles MB, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. sage; 1994.

36. Pekrun R, Frenzel AC, Goetz T, Perry RP. The control-value theory of achievement emotions: an integrative approach to emotions in education. In: *Emotion in Education*. Elsevier; 2007:13–36.

37. Hayat AA, Salehi A, Kojuri J. Medical student's academic performance: the role of academic emotions and motivation. *J Adv Med Educ Prof*. 2018;6(4):168.

38. Mega C, Ronconi L, De beni R. What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *J Educ Psychol*. 2014;106(1):121. doi:10.1037/a0033546

39. Reeve J. Autonomy-supportive teaching: what it is, how to do it. In: *Building Autonomous Learners*. Springer; 2016:129–152. doi:10.1007/978-981-287-630-0_7

40. Ten CO, Kusurkar RA, Williams GC. How self-determination theory can assist our understanding of the teaching and learning processes in medical education. AMEE Guide No., 59. *Med Teach*. 2011;33(12):961–973. doi:10.3109/0142159X.2011.595435

41. Das S, Nandi K, Baruah P, Sarkar SK, Goswami B, Koner BC. Is learning outcome after team based learning influenced by gender and academic standing? *Biochem Mol Biol Educ*. 2019;47(1):58–66. doi:10.1002/bmb.21197

42. Rotgans JI, Schmidt HG, Rajalingam P, et al. How cognitive engagement fluctuates during a team-based learning session and how it predicts academic achievement. *Adv Health Sci Educ*. 2018;23(2):339–351. doi:10.1007/s10459-017-9801-2

43. Greene BA, Miller RB, Crowson HM, Duke BL, Akey KL. Predicting high school students' cognitive engagement and achievement: contributions of classroom perceptions and motivation. *Contemp Educ Psychol*. 2004;29(4):462–482. doi:10.1016/j.cedpsych.2004.01.006

44. Teemant A, Hausman CS, Kigamwa JC. The effects of higher order thinking on student achievement and English proficiency. *INTESOL J*. 2016;13(1):51.

45. Hamzah H, Hamzah MI, Zulkifli H. Systematic Literature review on the elements of metacognition-based Higher Order Thinking Skills (HOTS) teaching and learning modules. *Sustainability*. 2022;14(2):813. doi:10.3390/su14020813

46. Sahoo S, Mohammed CA. Fostering critical thinking and collaborative learning skills among medical students through a research protocol writing activity in the curriculum. *Korean J Med Educ*. 2018;30(2):109. doi:10.3946/kjme.2018.86

47. Dubey S, Dubey AK. Promotion of higher order of cognition in undergraduate medical students using case-based approach. *J Educ Health Promot*. 2017;6. doi:10.4103/jehp.jehp_39_17

48. Roberson B, Franchini B. Effective task design for the TBL classroom. *J Excell Coll Teach*. 2014;25(3–4):303–333.

49. Lee J-Y, Lee D-Y. The effects of team-based learning program on the self-directed learning capability, learning transfer and learning satisfaction of nursing students. *Indian J Public Heal Res Dev*. 2018;9(11):1.

50. Knowles M. Self-Directed Learning. Englewood Cliffs, NJ: Cambridge Adult Education. 1975.

51. Kaufman DM. *Teaching and Learning in Medical Education: How Theory Can Inform Practice*. Swanwick T, K Forest, Bridget O, eds. 3rd ed. Wiley Online Library; 2018. doi: 10.1002/9781119373780.ch4

52. Murad MH, Coto-Yglesias F, Varkey P, Prokop LJ, Murad AL. The effectiveness of self-directed learning in health professions education: a systematic review. *Med Educ*. 2010;44(11):1057–1068. doi:10.1111/j.1365-2923.2010.03750.x

53. Janotha B. *The Effect of Team-Based Learning on Student Self-Directed Learning*. Hofstra University; 2015.

54. Hill M, Peters M, Salvaggio M, Vinnedge J, Darden A. Implementation and evaluation of a self-directed learning activity for first-year medical students. *Med Educ Online*. 2020;25(1):1717780. doi:10.1080/10872981.2020.1717780

55. Han SJ, Lee Y, Beyerlein M, Kolb J. Shared leadership in teams: the role of coordination, goal commitment, and knowledge sharing on perceived team performance. *Team Perform Manag an Int J*. 2018;24(4):150–168.

56. Dougherty RW, Wyles CC, Pawlina W, Lachman N. The team is more than the sum of its parts": implementation of charters to improve team dynamics in an anatomy course. *TAPS*. 2018;3(1):6–14. doi:10.29060/TAPS.2018-3-1/OA1050

57. Pentland AS. The new science of building great teams. *Harv Bus Rev*. 2012;90(4):60–69.

58. Tuckman BW, Jensen MAC. Stages of small-group development revisited. *Organ Stud*. 1977;2(4):419–427. doi:10.1177/105960117700200404

59. Elkhamisy FA, Wassef RM. Innovating pathology learning via Kahoot! Game-based tool: a quantitative study of students' perceptions and academic performance. *Alexandria J Med*. 2021;57(1):215–223. doi:10.1080/20905068.2021.1954413

60. Elkhamisy FA, Sharif AF. Medical students perceptions of virtual learning stations as an innovative teaching tool: a qualitative study. *Interact Learn Environ*. 2021;1–17. doi:10.1080/10494820.2021.2002366

61. Voigt C, Vogelsang K, Liere-Netheler K, Blömer L, Brink H, Hoppe U Understanding the Habits: inertia in Flipped Classroom. In: *International Conference on Business Informatics Research*. Austria: Springer; 2019:219–232.

62. Figueiró PS, Raufflet E. Sustainability in higher education: a systematic review with focus on management education. *J Clean Prod*. 2015;106:22–33. doi:10.1016/j.jclepro.2015.04.118

63. Chen O, Kalyuga S. Cognitive load theory, spacing effect, and working memory resources depletion: implications for instructional design. In: Hai-Jew S, editor. *Form, Function, and Style in Instructional Design: Emerging Research and Opportunities*. Educational Technologies and Instructional Design (AETID) Book Series; Vol. 1, 2020:1–4. doi:10.4018/978-1-5225-9833-6.ch001