

## Peer Assessment in the Context of Team-Based Learning in Undergraduate Education: How Far Can We Go?

Alba Regina de Abreu Lima<sup>1</sup>, Júlio César André<sup>2</sup>, Marina Kamimura<sup>3</sup>, Ana Carolina dos Reis<sup>4</sup>, Juanita Justina Ferreira da Silva<sup>4</sup>, Loiane Letícia dos Santos<sup>5</sup>, Emerson Roberto dos Santos<sup>6</sup>, Sérgio Luíz Aparecido Brienze<sup>7</sup>

<sup>1</sup>Faculty of Medicine of São José do Rio Preto - FAMERP, Fellow of the FAIMER Regional Institute (Foundation for Advancement of International Medical Education and Research)-BRAZIL 2015. (ORCID: <https://orcid.org/0000-0003-4332-4059>)

<sup>2</sup>Center for Studies and Development of Health Education, Faculty of Medicine of São José do Rio Preto-CEDES/FAMERP, Fellow of the FAIMER Regional Institute (Foundation for Advancement of International Medical Education and Research) - BRAZIL 2016. (ORCID: <https://orcid.org/0000-0002-0549-4527>)

<sup>3</sup>Undergraduate student in Medicine at the Faculty of Medicine of São José do Rio Preto - FAMERP, Scholarship in Programa Institucional de Bolsas de Iniciação Científica of FAMERP (PIBIC/FAMERP), Public notice GDG 029/2016 of April 1, 2016. (ORCID: <https://orcid.org/0000-0002-9008-8661>)

<sup>4</sup>Undergraduate students in Medicine at the Faculty of Medicine of São José do Rio Preto - FAMERP. (ORCID: <https://orcid.org/0000-0001-9044-1317>; <https://orcid.org/0000-0003-3172-0206>)

<sup>5</sup>Instituto Municipal de Ensino de Catanduva - IMES, Catanduva, SP. (ORCID: <https://orcid.org/0000-0003-0571-0370>)

<sup>6</sup>Master's student in Nursing of the Graduate Program in Nursing at the Faculty of Medicine of São José do Rio Preto - FAMERP, 2018. (ORCID: <https://orcid.org/0000-0002-9513-1083>)

<sup>7</sup>Faculty of Medicine of São José do Rio Preto - FAMERP, Fellow of the FAIMER Regional Institute (Foundation for Advancement of International Medical Education and Research)-BRAZIL 2017. (ORCID: <https://orcid.org/0000-0003-4765-2340>)

Correspondence: Alba Regina de Abreu Lima, Center for the Study and Development of Health Education – CEDES, Medical School of Medicine of São José do Rio Preto - FAMERP, Av. Brigadeiro Faria Lima, - 5416 - Vila São Pedro, Zip Code: 15090-000, São José do Rio Preto - SP Brazil.

Received: Dec. 3, 2019

Accepted: Feb. 23, 2020

Online Published: Feb. 27, 2020

doi:10.11114/jets.v8i3.4737

URL: <https://doi.org/10.11114/jets.v8i3.4737>

### Abstract

In medical education, the team-based learning method (TBL) is a teaching strategy used to intensify interactive learning in small groups, in which the student is given the role of evaluating his/her peers - peer assessment (PA). To investigate the interference of the students' interpersonal relationships in awarding their peers grades ("halo effect"). A qualitative and quantitative retrospective study. The study participants were 78 first-year medical students, divided into 17 teams for the TBL. The final grade of the PA for each member was calculated by the average of the grades received from their peers. Results: The comparison between the average of the evaluations in the TBL method (MTBLs) and the PA showed that 17.64% of the teams showed a significant difference between the grades, thus having the "halo effect". In the qualitative analysis, the "halo effect" was evidenced in only one of these teams. Although many studies corroborate the idea that using PA in the formative assessment is appropriate, advancing in the use of PA in the summative assessment is necessary, integrating it into the institution's evaluation system. Data presented here can help in continuing its use and in increasing its reliability.

**Keywords:** medical education, active methodologies, active learning, team-based learning, peer assessment

### 1. Introduction

Historically, the training of health professionals is based on positivist models using traditional methodologies (Giant & Fields, 2016). Current studies and assessments of medical schools in Brazil show most medical courses are still

organized according to the propositions of the Flexner report (Nogueira, 2009), published more than 100 years ago, in 1910. Thus, the training of medical professionals is still based on the mechanistic, uncritical, reproductive and individualistic health work (Giant & Fields, 2016).

However, medical education specialists have been challenged to create conditions to produce doctors who, in addition to being technically competent, have a critical, humanistic, ethical and reflective profile and are prepared to work collaboratively in a team, with social responsibility (Brasil, 2001; Brasil, 2014). Thus, to achieve this goal, changes in the teaching-learning method and environment have been encouraged by practical, political and pedagogical initiatives aimed at a more interactive, cooperative and significant process (Almeida 2001; Abrahao & Merhy, 2014).

In addition, in the recent “Knowledge Society”, the professionals are required to have new skills and competencies: teamwork, adaptability, search for creative and innovative solutions, application of knowledge and skills, openness to criticism, continuous updating through research, fluency in various languages, team management, computer and informatics knowledge and dialogue between peers. These demands cause higher education to think about innovation (Masetto, 2004).

### *1.1 Historical Context in Brazil*

Within this scenario, the active teaching and learning methodologies have emerged. They are understood as a model of professional training more compatible with the current health policy principles (Marin et al., 2010), as described in article 7 of the 2014 National Curriculum Guidelines of MEC, which provides that undergraduate students in Health education “should also take responsibility for their initial, continuous and in-service training, intellectual autonomy, bear their social responsibility in mind, as well as commit to the training of future generations of health professionals and encourage academic and professional mobility”. In the first section, we find the undergraduate student should “learn to learn, as part of the teaching-learning process, identifying previous knowledge, developing curiosity and framing questions for the search for scientifically consolidated answers, building meanings for professional identity and critically evaluating the information obtained, preserving the privacy of the sources”. In the same document, article 32 specifies that the undergraduate course in Medicine should use active methodologies and criteria for monitoring and assessment of the teaching-learning process and of the course itself, as well as develop tools to verify the structure, processes and outcomes, according to the National Higher Education Assessment System (SINAES) and the curricular dynamics defined by the Higher Education Institute, where it was implemented and developed (Brasil, 2007).

### *1.2 Related Literature*

The emergence of new strategies has been strengthening the study of active methodologies, seeking to promote the undergraduate student’s autonomy, both using simpler strategies and using those requiring a physical and/or technological readjustment of educational institutions (Farias et al., 2015). Activities and group works can be a highly effective way to master the basic subjects and develop problem-solving skills for both students and teachers (Frame et al., 2015).

The study conducted at Jefferson Medical College, in Pennsylvania, USA, showed active teaching-learning methodologies can directly benefit the student and the population. In this experience, students were divided into small groups and used strategies for discussing case studies. The results obtained by students at the National Board of Medical Examiners were compared with the students’ overall average, achieving better results since the first year of implementation of the active methodologies (Damjanov et al., 2005).

There are several modalities of active methodologies, and, to be considered good methodologies, they must be constructivist, based on meaningful learning; collaborative, favoring the construction of team knowledge; interdisciplinary; contextualized, allowing the students to understand the application of this knowledge in their reality; reflective, emphasizing moral and ethical principles; critical, encouraging the student to seek knowledge deepening; investigative, instigating the student's curiosity and autonomy; humanist, integrated into the social context; motivating and challenging, which encourages the student to find solutions (Cecy et al., 2010). In the Brazilian context is no different (Nogueira, 2009).

#### *1.2.1 Team-Based Learning (TBL)*

The so-called team-based learning, or TBL, is among the several active methodologies. In medical education, the team-based learning method (TBL) is a teaching strategy used in some universities to intensify interactive learning in small groups, mainly in the basic cycle. The TBL was first developed by Michaelsen, in the early 1990s, in response to four facts: increased number of students in his class, which went from 40 to 120 students; his own dissatisfaction with his classes and lectures; inability to know what and how their students thought during his classes; and the fact that students did not have opportunities to solve, in the classroom, problems they would have to solve in the real world (Parmelee et al., 2012).

In the TBL method, the class is divided into small groups, from 5 to 8 students, with as much heterogeneity as possible among its members. The team must have the same composition along the subject, and team members can evaluate their peers. The first stage is the pre-class preparation, which should be individual, including the reading of the recommended texts on the subject that will be dealt with in class and the analysis of the study material. The second step is the shared commitment, which is developed within the classroom. It starts with a quick individual test, and then the same test is done in groups, starting the group discussions. These discussions aim at exchanging experiences, so that the team reaches a consensus. After all the teams reach a consensus and solve the group tests, the whole class shares their answers, which then become the material for class discussion, in which the teacher reviews and explains the main teaching points. Finally, the third and final stage consists of applying the concepts of the subject. In this stage, a guided activity was done, observing the impact of the method on learning (Almeida, 2001; Ravindranath et al., 2010).

Forming groups makes the student learn to work as a team, to dialogue, to negotiate, and to contribute ideas (Brodbeck et al., 2002). In addition, the group work becomes a social demand, because if the student is unprepared at the time of discussion, the performance of his/her teammates will be affected, and he/she will be judged by them. With the TBL, the student stops being a passive receiver and becomes the main figure of his/her learning, self-managing his/her studies (Lerner & Tetlock, 1999).

### 1.2.2 Peer Assessment (PA)

In the TBL method, peer assessment can be performed. Traditionally, assessment in medical education is a responsibility essentially attributed to teachers (Vleuten & Schuwirth, 2005). However, students are increasingly held responsible for it, including taking the role of peer assessment agents, aiming at empowering doctors for their continued development and performance in teams (Vleuten & Schuwirth, 2005). According to Domingues (2007), "Peer assessment has been pointed out as a good indicator of future professional performance and is considered steady and reliable, providing information not measured by traditional methods" (Domingues et al., 2007).

Although promising and innovative, it is poorly used in medical education (Belar et al., 2001). This seems to be partly due to the reluctance of students to evaluate their peers and of universities to develop programs to prepare both students and professors to carry out conscious and constructive assessments, creating, in parallel, a safe and favorable assessment environment. A peer assessment enables students to actively participate in the knowledge construction. Moreover, it is indicated as a more effective method to empower the professional to work in a team and to be constantly developing; and students can improve their humanistic competence by analyzing the opinions of their peers (Domingues et al., 2007). In fact, peer assessment can be a powerful tool to assess and encourage the development of professional behaviors (Nofziger et al., 2010).

A study showed peer assessment can also improve the students' future professional behavior, finding that students who underwent peer assessment received better grades from their professors due to their professional behavior (Lurie et al., 2006). In another study conducted with classes submitted to peer assessment, 65% of the students described changes in consciousness, attitude and/or behavior due to the peer assessments. Most of these changes were described as positive and include speaking more in groups, increased patience and punctuality, or higher motivation (Nofziger et al., 2010).

However, few articles and studies address the introduction of the peer assessment method in the early stages of medical education in the undergraduate course (de Fátima Wardenski et al., 2012; Rezende et al., 2019). Thus, studying the impact and effectiveness of its introduction on this context is necessary.

### 1.2.3 "Halo Effect"

Evaluating whether the results of the peer assessment correspond to the students' theoretical performance in the TBL tests, and whether interpersonal relationships interfere in the grades the students award to their peers ("halo effect") is necessary. According to the Encyclopedia Britannica, the halo effect occurs when an impression is formed from an initial characteristic, influencing multiple judgments or classifications of unrelated factors. The pioneering study on the "halo effect" phenomenon was developed by the American psychologist Edward L. Thorndike, who "reported the occurrence of this effect on soldiers in 1920 through experiments in which the commanding officers were invited to evaluate their subordinates' intelligence, physical shape, leadership and personality, without having spoken to the subordinates" (Neugaard, 2016).

## 1.3 Purpose

To analyze the correlation between the grade received in the peer assessment, in the TBL test, and the final grade in the subject, this study investigates the reliability potential of the peer assessment in the TBL, to increase the application of the method, both in basic subjects and in other subjects, at different times of the course and in other courses.

## 2. Method

### 2.1 Research Context

Human Embryology, from the basic cycle of the medical course of the Faculty of Medicine of São José do Rio Preto – FAMERP, taught in the first year of the course. The subject of Human Embryology was chosen because it requires the students to quickly understand a set of changes occurring simultaneously on a macro and microscopic scale in the embryo and, consequently, students have difficulty in understanding the concepts presented and in mentally creating three-dimensional images of the processes involved. Also, it requires the teacher to be able to didactically describe complex and simultaneous processes. Therefore, the TBL may become a support tool.

In 2018, the subject of Human Embryology consisted of 16 subjects, 4 of which were presented using the TBL method; and the others, using lectures. The assessment in the subject was carried out through 3 partial tests (PTs), TBL tests (TBL), and peer assessment (PA). The PTs consisted of multiple-choice or essay tests administered at previously scheduled times while taking the subject. TBL tests consisted of multiple-choice tests according to the subject of the class to be taught and were first done individually (ITBL) and, then, in group (GTBL), at the end of each class using the TBL method.

### 2.2 Study Design

A qualitative and quantitative retrospective study. The students were randomly distributed in 17 teams varying from 3 to 5 students; thus, the sample was composed of 11 teams with 5 students, 5 with 4 students and only 1 team with 3 students (Figure 1).

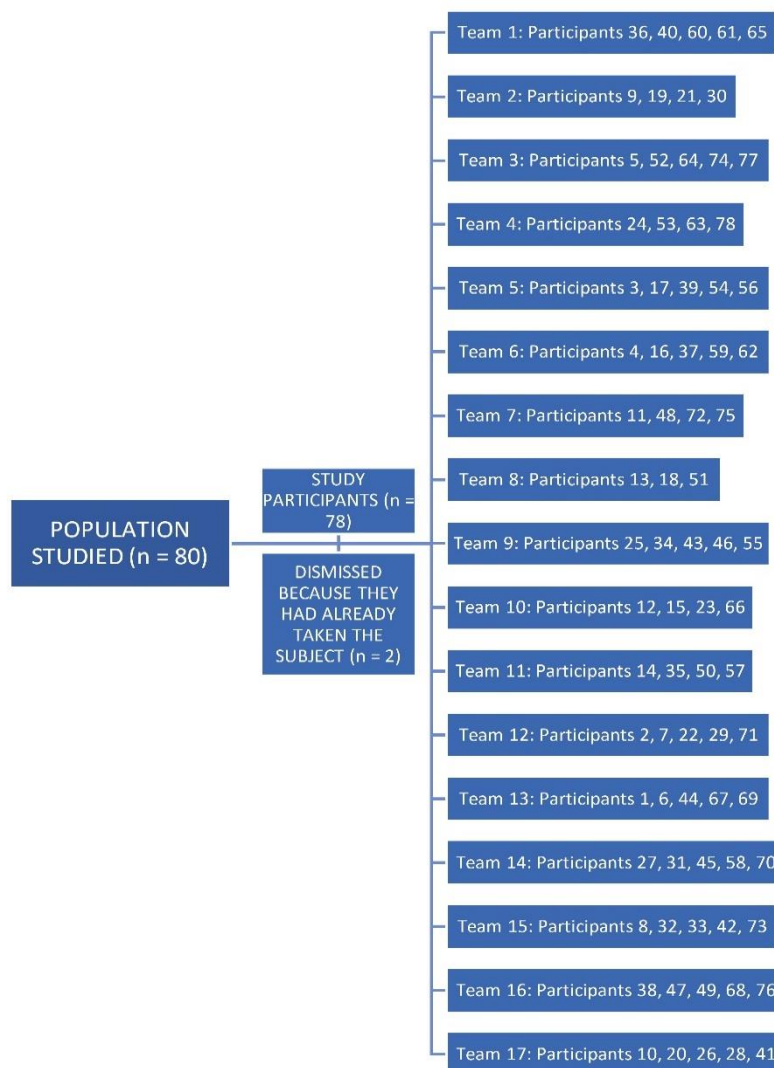


Figure 1. Sample characterization and random division of teams with randomly numbered students

### 2.3 Sampling

The population studied was a group of 80 students who are currently in the fifth year of the medical course of FAMERP and attended the subject of Human Embryology in 2018, that is, during the first year of the undergraduate course. The names of the students were concealed and they all received a random number to preserve their identity. The students of the medical course of FAMERP who were dismissed from taking the subject because they had already done it in other institutions were excluded from the study. Thus, the study participants were 78 out of 80 students, considering that 2 were dismissed because they had already taken the subject of Human Embriology in other institutions. Most of them were females (n = 42).

### 2.4 Data Collection

The information was collected in 2018 through the basic subject of Human Embryology.

#### 2.4.1 Quantitative

The PA was composed of the distribution of points reflecting how each student felt about the each of its team members contributed to their learning and/or to the team's performance, so that the students were asked to distribute from 0 to 15 points for each of the team members always differentiating the grades, that is, at least one member should receive a grade above 10 or higher (maximum 15), and another member should receive 10 or lower. The final grade of the PA for each member was calculated by the mean of the points received from their peers, the other team members (Figure 2).

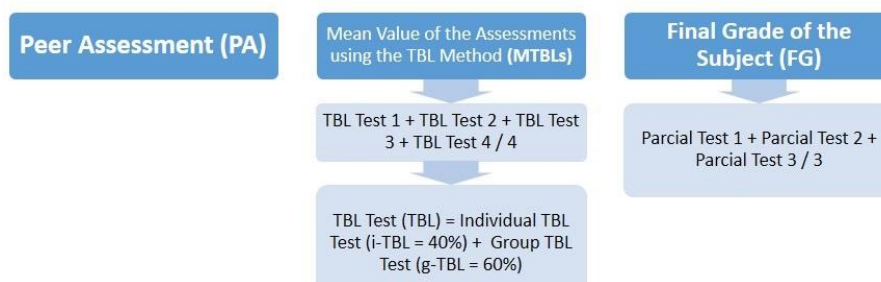


Figure 2. Examples of calculation of each student’s final grade in the peer assessment (PA)

#### 2.4.2 Qualitative

For the qualitative analysis, the grades attributed by the peers to each student and their justification for the assigned grades were analyzed. The inclusion of the justifications of the grades given by peers, an integral part of the model used here, is required in the process.

### 2.5 Data Analysis

#### 2.5.1 Quantitative

For each student, a PA, the mean value of four TBLs (MTBL). In each TBL, the individual TBL (i-TBL) represented 40% and group TBL (g-TBL) represented 60%, and the final grade of the subject (FG = simple arithmetic mean of the 3 PTs) (Figure 3).

3 PARTICIPANTS TEAMS		4 PARTICIPANTS TEAMS		5 PARTICIPANTS TEAMS	
MEMBERS	MEMBER 1 SCORES ASSIGNED (0 – 15)	MEMBERS	MEMBER 6 SCORES ASSIGNED (0 – 15)	MEMBERS	MEMBER 12 SCORES ASSIGNED (0 – 15)
Member 1	doesn't assign to himself	Member 4	11	Member 8	15
Member 2	15	Member 5	9	Member 9	8
Member 3	9	Member 6	doesn't assign to himself	Member 10	9
MEMBER 1 FINAL SCORE	12 POINTS (15 + 9 / 2)	Member 7	12	Member 11	10
		MEMBER 6 FINAL SCORE	10.7 POINTS (11 + 9 + 12 / 3)	Member 12	doesn't assign to himself
				MEMBER 12 FINAL SCORE	10.5 POINTS (15 + 8 + 9 + 10 / 4)

Figure 3. Diagram of the data analyzed for each student

All data were converted on the same scale from 0 to 10 and tabulated. Subsequently, normality and dispersion tests were performed, and the parametric analysis of variance (ANOVA) was used to compare the grades achieved in the PA, the MTBLs, and the FG.

### 2.5.2 Qualitative

In cases in which the result showed the distribution of some teams differed from the others, but did not indicate between which teams the difference was significant, the Tukey's multiple comparison test was used (Risucci et al., 1992).

### 2.6 Ethical Aspects

The Research Ethics Committee of the Faculty of Medicine of São José do Rio Preto (CEP – FAMERP) approved this project (CAAE: 28979514.4.000.5415).

## 3. Results

Statistically significant differences (ANOVA test and post hoc Tukey) were found between the PA grades and the MTBLs in teams 9 ( $P=0.0172$ , and post hoc Tukey  $P<0.05$ ), 14 ( $P=0.0448$ , and post hoc Tukey  $P<0.05$ ) and 17 ( $P=0.0073$ , and post hoc Tukey  $P<0.05$ ). In addition, a statistically significant difference was obtained between the MTBLs and the FG of the subject in teams 6 ( $P=0.0302$ , and post hoc Tukey  $P<0.05$ ), 16 ( $P=0.0207$ , and post hoc Tukey  $P<0.05$ ), and 17 ( $P=0.0073$ , and post hoc Tukey  $P<0.05$ ).

In teams 6 ( $P=0.0302$ ), 16 ( $P=0.0207$ ), and 17 ( $P=0.0073$ ), in which we found a statistically significant difference between the MTBLs and the FG, the MF was lower than the MTBLs (Table 1).

Table 1. Results of the ANOVA test

Team	ANOVA	PA x MTBLs	MTBLs x FG
1	$P=0.0988$		
2	$P=0.2130$		
3	$P=0.0741$		
4	$P=0.3206$		
5	$P=0.1953$		
6	$P=0.0302$		POST HOC TUKEY $P<0.05$
7	$P=0.8842$		
8	$P=0.6364$		
9	$P=0.0172$	POST HOC TUKEY $P<0.05$	
10	$P=0.0901$		
11	$P=0.1979$		
12	$P=0.2679$		
13	$P=0.6567$		
14	$P=0.0448$	POST HOC TUKEY $P<0.05$	
15	$P=0.9262$		
16	$P=0.0207$		POST HOC TUKEY $P<0.05$
17	$P=0.0073$	POTS HOC TUKEY $P<0.05$	POST HOC TUKEY $P<0.05$

PA = Peer assessment; MTBLs = mean value of the assessments using the TBL method; FG = Final grade of the subject)

In teams 9 ( $P=0.0172$ ), 14 ( $P=0.0448$ ), and 17 ( $P=0.0073$ ), in which a statistically significant difference was observed between the average of the PA grades and the MTBLs, the “halo effect” occurred.

In a qualitative analysis, the halo effect occurred in the team 17. When analyzing the grades awarded from each student to their peers, two members awarded the highest grade to the member 28; however, he received the lowest grade from the member 10. The former, when justifying the highest grades they awarded, affirmed that the member 28 was more collaborative regarding the resolution of issues and doubts, participated the most in the discussions and seemed to know the subject, while the member 10 stated that the participation of individual 28 was equal to that of the other members (Table 2).

In addition, the member 10 was the only one who awarded the highest grade to the member 41, stating that he was the most devoted and studied harder, contradicting the other members, who gave lower grades to the member 41.

Table 2. Justifications for grades in Team 17

PARTICIPANT	JUSTIFICATIONS FOR HIS HIGHEST GRADE	JUSTIFICATIONS FOR HIS LOWER GRADE
10	"The highest score was for the person I realized who studied more and worked harder." (highest score was given to participant 41)	"He was given the lowest score because it couldn't be equal, but I think he worked and helped as much as the others." (his lowest score was given to participant 26)
20	NOT JUSTIFICATIVE (his highest score was given to participant 28)	NOT JUSTIFICATIVE (his lowest score was given to participant 10)
26	"Everyone contributed a lot, but [participant 26 and participant 20] were the ones who discussed the most and seemed to be more aware of the subject." (their highest scores were given to participants 28 and 20)	"The lowest scores were only due to his participation, but at the time of the discussion, everyone worked well." (their lowest scores were attributed to participants 41 and 20)
28	"Always proactive, bringing content to the discussion and willing to talk about it." (highest score was given to participant 20)	"Little participative, however, this is probably due to the group's performance and the fact that the person has already been shy at our tests, which may have made it difficult to participate." (lowest score was given to participant 41)
41	"People collaborated more with the resolution of questions and doubts that arose." (their highest scores were attributed to participants 26 and 20)	"There was little collaboration in discussions on the issues." (lowest score was given to participant 10)

In the other two groups whose results were significant between the average of the PA grades and the MTBL, no data explicitly show the presence of "halo effect", when analyzing the grades and justifications from each member. The grades given by team members were similar.

Thus, in this study, regarding the comparison between the MTBL and the FG, only 17.64% of the teams (3 out of 17) showed a significant difference between the grades. Likewise, only 17.64% of the teams (3 out of 17) showed a significant difference between the average of the grades in the MTBL and in the PA. In the qualitative analysis, the "halo effect" was evidenced in only one of these teams.

In one of the teams, a student disagreed with the others when awarding grades, which favored one participant and disadvantaged another one. However, in his justification, no reason was found that explains this disparity, which points that the grade was given according to his personal affinities.

When analyzing the other two teams, no disagreement is observed between the members when awarding grades to their teammates, that is, the best and worst grades were always given to the same members. Thus, one can assume that the "halo effect", in these teams, results from a worse performance in the individual TBL tests, and not from favoritism due to interpersonal relationships between the members.

Although the study sample in question is small, in most teams (82.36%, 14 out of 17), the students' individual performance as well as their knowledge demonstrated during the discussions was more valued than their relationships with other team members.

#### 4. Discussion

TBL has proven efficacy in different undergraduate courses in health sciences such as Nursing (Reimschisel et al., 2017; Dearnley et al., 2018), Pharmacy (Basheti et al., 2010; Dearnley et al., 2018), Physiotherapy (Maas et al., 2014), Speech Therapy (Nascimento et al., 2019), Medicine (Garner et al., 2010; Reimschisel et al., 2017; Cunha et al., 2019), and many others (Dearnley et al., 2018), at different periods of the course, since the first year (Santana et al., 2019), as in the case of this study, in intermediate (Masocatto et al., 2019) and final (Reimschisel et al., 2017) moments, on different continents (Reimschisel et al., 2017). It is also successfully at graduate level (Silva et al., 2017), including in medical residency (Kisiel et al., 2010; Thomas et al., 2011; McMullen et al., 2014).

Unlike cases in which the teacher works simultaneously with an entire class, in this case, with 78 students (Figure 1), for being subdivided into 3 from 5 teams and working together for several sessions (4 sessions), students are able to provide more accurate assessments of skills such as teamwork, communication, and professionalism (Papinczak et al., 2007; Epstein, 2007).

PA has been used in different instructional methods and at different levels. It is particularly useful in learning processes in which students work together towards a common goal, such as in the TBL and the problem-based learning (PBL). In the evaluation in tutorials in PBL, PA proved to be a valuable opportunity, in which some students and some groups of students were able to accurately judge the performance of their classmates in PBL tutorials (Papinczak et al., 2007), and its use for the evaluation of professional competencies and skills in residents and professional colleagues has been central in the referral process in medicine (Norcini, 2003).

Assuming that PA can help in identifying students who may have problems related to professionalism (Emke et al., 2015), its use should be considered to identify learning problems, which was not done in this study. It is known that PA can benefit students in their learning process, since it increases their analytical skills as well as their ability to achieve their learning goals and fulfill tasks related to problem analysis (Tayem et al., 2015).

One should consider that implementing PA in any scenario initially may lead to skepticism and doubts about its value and trust and implementing PA without preparing students properly can create an undesirable environment in class, which includes mistrust, increased competitiveness or even tendency to strive less than when working alone (Cestone et al., 2008; Levine, 2008). This preparation, or "education", for PR should highlight its importance and usefulness, mainly by modeling teaching roles (Arnold et al., 2005).

A PA system whose characteristics meet participants' concerns considering the main discouraging factors for adherence and honesty in the process (intimate personal conflicts, perspective of breach of anonymity and confidentiality, factors related to the educational environment, among others - Arnold et al., 2005) could minimize the reluctance and anxiety inherent in the process.

Qualitative results of studies conducted two decades ago had already demonstrated the overestimation of peer performance resulting from personal relationships such as friendship, which leads to biased answers in the PA (Pond & ul-Haq, 1997). Some students strongly believe that colleagues may be overly condescending, not making an honest assessment (Cestone et al., 2008). One must consider that the absence of preparation of students for PA may have contributed to the occurrence of the "halo" effect reported herein.

The "halo effect" may be the reluctance of students to eventually make their peers fail the subject, as reported by Epstein (Epstein, 2007). Thus, using PA to provide feedback on professional behaviors such as work habits, interpersonal behavior, and team skills is preferable. This limitation is well known and expected, which limits the use of PA to formative and summative assessment. In the TBL, the peer scoring rubric and the method used reduce this effect, but without eliminating it. Although many studies corroborate the idea that using PA in the formative assessment is appropriate, advancing in its use in the summative assessment is necessary, an issue also pointed out in the literature (Tayem et al., 2015). This is because PA can be feasible and useful, when the steps that ensure its success and reliability are taken, including the training of teachers and students on its methods and purposes and the creation of a collaborative learning environment. In a collaborative learning environment where students are likely to teach one another and explore issues of professionalism, PA contributes to a collaborative and productive criticism of peer performance, in contrast to groups in which the competition, protectionism and trench mindset discourage honest assessment (Arnold et al., 2005).

#### *4.1 Implications of the Findings*

This study was conducted within a less updated curriculum framework (serial, annual, subject-based, and divided into basic, clinical and internship cycle). However, the use of TBL alone, in an activity (4 in this study) or during the course, is known to contribute to a movement towards reflection and action for a more active learning in the health courses, since it is a pedagogical strategy based on key principles of adult learning (Holterman, 1999). This strategy values the individual responsibility of students in their work teams and has a motivational component for the study, which is the application of the knowledge acquired for resolution of relevant issues in the professional practice (Bollela et al., 2014).

Although the FG was lower than that of the MTBLs in only 3 out of 17 teams (Table 1), one should consider that using the test can positively influence the learning outcomes, since any new instructional method has, as a rule, the benefit of awakening the highest level of interest, which can translate into better academic results, perhaps more associated with the fact that it is a new and more attractive modality than the modality itself (Rezende et al., 2019).

The inclusion of the justifications of the grades given by peers, an integral part of the model used here, is required in the process; otherwise, it would mischaracterize and contradict the fact that "narrative comments provide, in addition to data quantitative evaluation, valuable information on interpersonal dynamics" (Kisiel et al., 2010), since all information collected in the PA, if used properly, can ensure more discriminatory feedback, collaborating to overcome difficulties and finally helping those being evaluated become better doctors (Thomas et al., 2011).

#### *4.2 Limitations of the Study*

We compared the results of the PA in the TBL method using only one instructional method, formal expository classes. Comparing the results of the PA in the TBL with the results of other instructional methods that are equally active is necessary.

Only one class, in a single moment it has been studied, needed follow-up in other moments during the course to investigate the behavior of these differences and of the "halo effect".



In this study, students were not prepared for PA. “Education” towards the performance of peer assessment was not provided, which may have caused the “halo effect”. Awareness-raising tests regarding their readiness were not performed as well. In a recent systematic review, out of 31 studies, 11 emphasized the importance of preparing students for PA, observing that they may feel confused or not know how to evaluate their peers properly without orientations or training. A more comprehensive instruction on PA can make its goals clearer, as well as decrease students' anxiety about evaluating their peers (Lerchenfeldt et al., 2019), but the anonymity in PA, recommended (Basheti et al., 2010), and here strongly preserved in this study, collaborates to avoid discomfort and possible bias in grades due to interpersonal relationships. However, if anonymity can protect the evaluator and the peer to be evaluated, some students believe anonymous assessments do not provide them with an adequate preparation for a future task of confronting colleagues face to face at work (Arnold et al., 2005).

#### 4.3 Future Suggested Studies

Considering the spread of the use of TBL and PA, we believe data from this study, particularly when it comes to continuing the use and increasing the reliability of PA, is a universal appeal. For this purpose, further studies including the feedback provided in the PA and the effect on academic performance, on institutional culture and on benefits for future employers and patients are required. As only one PA was performed at the end of the subject, we did not evaluate whether peer assessment help in spotting students with learning difficulties, on time to change our plans, which is highly recommended.

The use of anonymous assessments in this model of PA should be followed by the use of non-anonymous assessments in order to contribute to an adequate preparation for the future task of confronting colleagues face to face at work. The students themselves agree, although partially, with the use of PA in the summative assessment, for example, when a peer does not change his/her unprofessional behavior despite previous warnings or when the peer's mistake is very serious (Arnold et al., 2005).

#### Acknowledgements

The authors gratefully acknowledge PhD. Lilian Castiglioni's statistical advice.

#### References

- Abraham, A. L., & Merhy, E. E. (2014). Health education and micropolitics: about tool concepts in teaching practice. *Interface (Botucatu, Online)*, 18(49), 313-324. <https://doi.org/10.1590/1807-57622013.0166>
- Almeida, M. J. (2001). Medical education and the current proposals for change: some historical background. *Revista Brasileira de Educação Médica*, 25(2), 17-33.
- Arnold, L., Shue, C. K., Kritt, B., Ginsburg, S., & Stern, D. T. (2005). Medical Students' Views on Peer Assessment of Professionalism. *Journal of General Internal Medicine*, 20(9), 819-824. <https://doi.org/10.1111/j.1525-1497.2005.0162.x>
- Basheti, I. A., Ryan, G., Woulfe, J., & Bartimote-Aufflick, K. (2010). Anonymous peer assessment of medication management reviews. *American Journal of Pharmaceutical Education*, 74(5), 77. <https://doi.org/10.5688/aj740577>
- Belar, C., Brown, R. A., Hersch, L. E., Hornyak, L. M., Rozensky, R. H., Sheridan, E. P., ... Reed, G. W. (2001). Self-assessment in clinical health psychology: a model for ethical expansion of practice. *Professional Psychology, Research and Practice*, 32(2), 135-141. <https://doi.org/10.1037/0735-7028.32.2.135>
- Bollela, V. R., Senger, M. H., Tourinho, F. S. V., & Amaral, E. (2014). Team-based learning: from theory to practice. *Medicina (Ribeirão Preto)*, 47(3), 293-300. <https://doi.org/10.11606/issn.2176-7262.v47i3p293-300>
- Brasil. Ministério da Educação (2001). *National Curriculum Guidelines for Undergraduate Nursing, Medicine and Nutrition Courses*. Retrieved from <http://portal.mec.gov.br/dmdocuments/ces1133.pdf>
- Brasil. Ministério da Educação. (2007). *Opinion CNE / CES 116/2014. Institutes National Curriculum Guidelines for Undergraduate Medicine*. Retrieved from [http://portal.mec.gov.br/index.php?option=com\\_docman&view=download&alias=15514-pces116-14&category\\_slug=abril-2014-pdf&Itemid=30192](http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=15514-pces116-14&category_slug=abril-2014-pdf&Itemid=30192)
- Brasil. Ministério da Educação. (2014). *Diretrizes Curriculares Nacionais do Curso de Graduação em Medicina*. Retrieved from <http://portal.mec.gov.br/cne/arquivos/pdf/Med.pdf>
- Brodbeck, F. C., Kerschreiter, R., Mojzisch, A., Frey, D., & Hardt, S. S. (2002). The dissemination of critical, unshared information in decision-making groups: the effects of pre-discussion dissent. *European Journal of Social Psychology*, 32(1), 35-56. <https://doi.org/10.1002/ejsp.74>
- Cecy, C., Oliveira, G.A., & Costa, E. M. M. B. (2010). *Active methodologies: applications and experiences in*

- pharmaceutical education*. Brasília, DF, Associação Brasileira de Ensino Farmacêutico e Bioquímico.
- Cestone, C. M., Levine, R. E., & Lane, D. R. (2008). Peer assessment and evaluation in team-based learning. *New Directions for Teaching & Learning*, 116, 69-78. <https://doi.org/10.1002/tl.334>
- Cunha, C. R. O. B. J., Ramsdorf, B. M., Bragato, S. G. R. (2019). The Use Of Team-Based Learning As An Evaluation Method On The Medicine Course. *Revista Brasileira de Educação Médica*, 43(2), 208-215. <https://doi.org/10.1590/1981-52712015v43n2rb20180063>
- Damjanov, I., Fenderson, B. A., Hojat, M., & Rubin, E. (2005). Curricular Reform May Improve Students' Performance on Externally Administered Comprehensive Examinations. *Croatian Medical Journal*, 46(3), 443-448. Retrieved from <http://www.cmj.hr/2005/46/3/15861525.pdf>
- de Fátima Wardenski, R., de Espíndola, M. B., Struchiner, M., & Giannella, T. R. (2012). Blended learning in biochemistry education: analysis of medical students' perceptions. *Biochemistry and Molecular Biology Education*, 40(4), 222-228. doi: <https://doi.org/10.1002/bmb.20618>
- Dearnley, C., Rhodes, C., Roberts, P., Williams, P., & Prenton, S. (2018). Team based learning in nursing and midwifery higher education: a systematic review of the evidence for change. *Nurse Education Today*, 60, 75-83. <https://doi.org/10.1016/j.nedt.2017.09.012>
- Domingues, R. C. L., Amaral, E., & Zeferino, A. M. B. (2007). Peer self-assessment and evaluation: strategies for the professional development of the physician. *Revista Brasileira de Educação Médica*, 31(2), 173-175. <https://doi.org/10.1590/S0100-55022007000200008>
- Emke, A. R., Cheng, S., Dufault, C., Cianciolo, A. T., Musick, D., Richards, B., & Violato C. (2015). Developing Professionalism via Multisource Feedback in Team-Based Learning. *Teaching and Learning in Medicine*, 27(4), 362-365. <https://doi.org/10.1080/10401334.2015.1077135>
- Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine*, 356(4), 387-396. <https://doi.org/10.1056/NEJMra054784>
- Farias, P. A. M., Martin, A. L. A. R., & Cristo, C. S. (2015). Active learning in health education: historical background and applications. *Revista Brasileira de Educação Médica*, 39(1), 143-150. <https://doi.org/10.1590/1981-52712015v39n1e00602014>
- Frame, T. R., Cailor, S. M., Gryka R. J., Chen, A. M., Kiersma, M. E., & Sheppard, L. (2015). Student perceptions of team-based learning vs traditional lecture-based learning. *American Journal of Pharmaceutical Education*, 79(4), 51. <https://doi.org/10.5688/ajpe79451>
- Garner, J., McKendree, J., O'Sullivan, H., & Taylor, D. (2010). Undergraduate medical student attitudes to the peer assessment of professional behaviours in two medical schools. *Education for Primary Care*, 21(1), 32-37. <https://doi.org/10.1080/14739879.2010.11493873>
- Giant, R. L., & Fields, G. W. S. (2016). Continuing training and education in health policy in brazil: legal basis and theoretical references. *Trabalho, Educação e Saúde*, 14(3), 747-763. <https://doi.org/10.1590/1981-7746-sip00124>
- Holterman, M. J. (1999). Clinically relevant embryology. New approaches to education. *Pediatrics*, 104(3), 784. Retrieved from [https://pediatrics.aappublications.org/content/pediatrics/104/Supplement\\_4/762.full.pdf](https://pediatrics.aappublications.org/content/pediatrics/104/Supplement_4/762.full.pdf)
- Kisiel, J. B., Bundrick, J. B., & Beckman, T. J. (2010). Resident physicians' perspectives on effective outpatient teaching: A qualitative study. *Advances in Health Sciences Education: Theory and Practice*, 15(3), 357-368. <https://doi.org/10.1007/s10459-009-9202-2>
- Lerchenfeldt, S., Mi, M., & Eng, M. (2019). The utilization of peer feedback during collaborative learning in undergraduate medical education: a systematic review. *BMC Medical Education*, 19(1), 321. <https://doi.org/10.1186/s12909-019-1755-z>
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of Accountability. *Psychological Bulletin*, 125(2), 255-275. <https://doi.org/10.1037/0033-2909.125.2.255>
- Levine, R. E. (2008). Peer Evaluation in Team-Based Learning. In: L. K. Michaelsen, D. X. Parmalee, K. K. Mc Mahon, & R. E. Levine, (Eds.), *Team-Based Learning for Health Professions Education: A Guide to Using Small Groups for Improving Learning* (pp. 103-116). Sterling, VA: Stylus.
- Lurie, S. J., Nofziger, A. C., Meldrum, S., Mooney, C., & Epstein, R. M. (2006). Effects of rater selection on peer assessment among medical students. *Medical Education*, 40(11), 1088-1097. <https://doi.org/10.1111/j.1365-2929.2006.02613.x>

- Maas, M. J., Sluijsmans, D. M., van der Wees, P. J., Heerkens, Y. F., Nijhuis-van der Sanden, M. W., & van der Vleuten, C. P. (2014). Why peer assessment helps to improve clinical performance in undergraduate physical therapy education: A mixed methods design. *BMC Medical Education*, *14*, 117. <https://doi.org/10.1186/1472-6920-14-117>
- Marin, M. J. S., Gomes, R., Marvulo, M. M. L., Primo, E. M., Barbosa, P. M. K., & Druzian, S. (2010). Multiprofessional health-related graduate courses: results from experiences using active methodologies. *Interface (Botucatu, Online)*, *14*(33), 331-344. <https://doi.org/10.1590/S1414-32832010000200008>
- Masetto, M. (2004). Innovation in Higher Education. *Interface (Botucatu, Online)*, *8*(14), 197-202. <https://doi.org/10.1590/S1414-32832004000100018>
- Masocatto, N. O., Couto, J. W., Matta, T. S., & Porfirio, J. G. M. (2019). Undergraduate Medical Students' Views on Team-Based Learning (TBL). *Revista Brasileira de Educação Médica*, *43*(3), 111-114. <https://doi.org/10.1590/1981-52712015v43n3rb20180231>
- McMullen, I., Cartledge, J., Finch, E., Levine, R., & Iversen A. (2014). How we implemented team-based learning for postgraduate doctors. *Medical Teacher*, *36*(3), 191-195. <https://doi.org/10.3109/0142159X.2014.875617>
- Nascimento, E. N., Moreira, G. M. O., Tolentino, K. G. L., & Frizzo, A. C. F. (2019). Team Based Learning in Speech, Language and Hearing Sciences: experience in the Public Health qualification. *Revista CEFAC*, *21*(2), e15918. <https://doi.org/10.1590/1982-0216/201921215918>
- Neugaard, B. (2016). *Halo Effect*. Retrieved from <https://www.britannica.com/topic/halo-effect>
- Nofziger, A. C., Naumburg, E. H., Davis, B. J., Mooney, C. J., & Epstein, R. M. (2010). Impact of Peer Assessment on the Professional Development of Medical Students: A Qualitative Study. *Academic Medicine*, *85*(1), 140-147. <https://doi.org/10.1097/ACM.0b013e3181c47a5b>
- Nogueira, M. I. (2009). Changes in brazilian medical education in perspective: reflections on the emergence of a new way of thinking. *Revista Brasileira de Educação Médica*, *33*(2), 262-70. <https://doi.org/10.1590/S0100-55022009000200014>
- Norcini, J. (2003). Peer assessment of competence. *Medical Education*, *37*(6), 539-543. <https://doi.org/10.1046/j.1365-2923.2003.01536.x>
- Papinczak, T., Young, L., Groves, M., & Haynes, M. (2007). An analysis of peer, self, and tutor assessment in problem-based learning tutorials. *Medical Teacher*, *29*(5), e122-e132. <https://doi.org/10.1080/01421590701294323>
- Parmelee, D., Michaelsen, L. K., Cook, S., & Hudes, P. D. (2012). Team-based learning: A practical guide: AMEE Guide. *Medical Teacher*, *34*(65), 275-287. <https://doi.org/10.3109/0142159X.2012.651179>
- Pond, K., & ul-Haq, R. (1997). Learning to assess students using peer review. *Studies in Educational Evaluation*, *23*(4), 331-348. [https://doi.org/10.1016/S0191-491X\(97\)86214-1](https://doi.org/10.1016/S0191-491X(97)86214-1)
- Ravindranath, D., Gay, T. L., & Riba, M. B. (2010). Trainees as teachers in team-based learning. *Academic Psychiatry*, *34*(4), 294-297. <http://dx.doi.org/10.1176/appi.ap.34.4.294>
- Reimschisel, T., Herringb, A. L., Huang, J., & Minor, T. J. (2017). A systematic review of the published literature on team-based learning in health professions education. *Medical Teacher*, *39*(12) 1227-1237. <https://doi.org/10.1080/0142159X.2017.1340636>
- Rezende, A. B., de Oliveira, A. G. F., Vale, T. C., Teixeira, L. A. S., Lima, A. R. A., Lucchetti, A. L. G., ... Ezequiel, O. S. (2019). Comparison of Team-Based Learning versus Traditional Lectures in Neuroanatomy: Medical Student Knowledge and Satisfaction. *Anatomical Sciences Education*, *0*, 1-11. <https://doi.org/10.1002/ase.1926>
- Risucci, D. A., Lutsky, L., Rosati, R. J., & Tortolani, A. J. (1992). Reliability and accuracy of resident evaluations of surgical faculty. *Evaluation & the Health Professions*, *15*(3), 313-324. doi: <https://doi.org/10.1177/016327879201500304>
- Santana, V. C., Oliveira, C. R., & Ramos, R. B. (2019). First-year students' perceptions of team-based learning in a new medical genetics course. *Revista Brasileira de Educação Médica*, *43*(3), 170-177. <https://doi.org/10.1590/1981-52712015v43n3rb20180256>
- Silva Junior, G. B., Medeiros, M. A. S., Oliveira, J. G. R., Catrib, A. M. F., & Jardim, M. H. (2017). Team-Based Learning: Successful Experience in a Public Health Graduate Program. *Revista Brasileira de Educação Médica*, *41*(3), 397-401. <https://doi.org/10.1590/1981-52712015v41n3rb20160081>
- Tayem, Y. I., James, H., Al-Khaja, K. A. J., Razzak, R. L. A., Potu, B. K., & Sequeira, R. P. (2015). Medical students' perceptions of peer assessment in a problem-based learning curriculum. *Sultan Qaboos University Medical Journal*,

15(3), e376-e381e. <https://doi.org/10.18295/squmj.2015.15.03.012>

Thomas, M. R., Beckman, T. J., Mauck, K. F., Cha, S. S., & Thomas, K. G. (2011). Group assessments of resident physicians improve reliability and decrease halo error. *Journal of General Internal Medicine*, 26(7), 759-764. <https://doi.org/10.1007/s11606-011-1670-4>

van der Vleuten, C. P., & Schuwirth, L. W. (2005). Assessing professional competence: from methods to programmes. *Medical Education*, 39(3), 309-317. <https://doi.org/10.1111/j.1365-2929.2005.02094.x>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the [Creative Commons Attribution license](#) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.