Peer Instruction-Based Feedback Sessions Improve the Retention of Knowledge in Medical Students

Feedback Baseado em “Peer Instruction” Melhora a Retenção de Conhecimento em Estudantes de Medicina

Silvia Maria Riceto Ronchim Passeri
Eric Mazur

ABSTRACT
Peer Instruction (PI) is an interactive teaching-learning process between colleagues and has been applied in various universities throughout the world. This active teaching methodology improves students’ performance and their capacity to resolve problems when they perform activities with their study colleagues. There are no systematic studies about the use of PI in assessment feedback. The aim of our study is to identify whether the use of PI on assessment feedback improves the retention of basic concepts in medical programs. For this study 226 undergraduate students (Y2 = 115, Y3 = 111) enrolled in a Brazilian medical school were invited to participate. After taking the regular exam (RE), the students of the control group (125) could individually receive feedback (review of the exam) from the professor according to the course routine, and the students in the study group (101) were invited to participate in an immediate intervention after the RE with a feedback developed session using the peer instruction teaching method. At the conclusion of the feedback session, the students again answered the post-feedback exam (PFE) so that we could identify any changes in the answers compared with the regular exam taken before feedback and 6 months later, we applied a diagnostic exam (DE) of identify whether the students retained the concepts covered in the previous exams. The control and study groups are statistically significantly different in the RE (p = 0.0014) and DE (p < 0.000). The study group demonstrated better performance in both exams than the control group. When we gave feedback, using PI immediately after the exam, retention of basic science knowledge jumped to 39%, increasing by 15%. The students that had assessment feedback had the opportunity to discuss their misconceptions. These students had the highest number of correct answers with assimilate knowledge and fewer assimilation of wrong answers, therefore, students who received immediate feedback had less tendency to make the same conceptual errors. PI in the feedback was effective in improving retention of basic science knowledge.
PALAVRAS-CHAVE

– Instrução pelos colegas.
– Educação Médica.
– Feedback.

RESUMO

Instrução de pares (PI) é um processo interativo de ensino-aprendizagem entre os estudantes e tem sido aplicado em várias universidades em todo o mundo. Essa metodologia de ensino ativa melhora o desempenho dos alunos e sua capacidade de resolver problemas quando realizam atividades com seus colegas. Não há estudos sistemáticos sobre o uso de PI no feedback da avaliação. O objetivo do nosso estudo é identificar se o uso do PI no feedback da avaliação melhora a retenção de conceitos básicos em educação médica. Foram convidados a participar deste estudo 226 estudantes de graduação (segundo ano = 115, terceiro ano = 111) matriculados em uma escola de Medicina no Brasil. Após o exame regular (ER), os alunos do grupo controle (125) poderiam solicitar a revisão de prova de acordo com a rotina do curso, e os alunos do grupo de estudo (101) foram convidados a participar de uma intervenção imediata após o ER com uma sessão de feedback usando-se o método de ensino por instrução de pares. No final do feedback, os alunos responderam novamente ao exame pós-feedback (PFE) para que pudéssemos identificar quaisquer alterações nas respostas em comparação ao exame regular feito antes do feedback. Após seis meses, aplicamos um exame de diagnóstico (DE) para identificar se os alunos mantiveram os conceitos abordados nos exames anteriores. O desempenho dos estudantes dos grupos controle e estudo são estatisticamente diferentes no RE (p = 0,0014) e no DE (p < 0,000). O grupo de estudo demonstrou melhor desempenho em ambos os exames do que o grupo controle. Com a sessão de feedback, usando-se PI imediatamente após o exame, a retenção do conhecimento básico foi de 39%, aumentando em 15%. Os alunos que tiveram feedback de avaliação tiveram a oportunidade de discutir suas dificuldades. Esses alunos apresentaram o maior número de respostas corretas assimiladas e menor assimilação de respostas erradas. Portanto, os alunos que receberam feedback imediato apresentaram menor tendência a cometer os mesmos erros conceituais da primeira avaliação. PI no feedback foi eficaz em melhorar a retenção de conhecimentos básicos em estudantes de Medicina.

INTRODUCTION

With the current emphasis on quality improvement and accompanying cost containment, it is incumbent on medical universities to lead the way towards the most effective teaching strategies. While it is necessary to be aware of the broad scope of competencies required in the medical curriculum, it is also necessary to involve the students in the teaching-learning process to increase their effectiveness and their ability to use quality assessment tools, even before these students enter clinical practice.

There is no standard protocol for assessing a medical student’s progress and frequently the assessment of learning is limited to only fulfilling the norms of the Institution in classifying which students may proceed with their studies, and which must repeat a course because they have not attained a pre-established performance goal. Few courses include the assessment of a learning instrument or promote continuous improvement of the course and of the teaching-learning process itself. Courses that offer feedback on the assessment are rare; that is, in many cases the student does not identify his/her difficulties, and does not learn from his/her errors. Feedback may help students attain their full potential and scientific evidence has proven the power of feedback in learning.

Some instructors believe that they provide their students with feedback; however, many students say they receive no feedback. Clearly this process is not efficient, particularly as many instructors receive no training on how to provide feedback, and we now have a generation of students who may not value critical reflection of the teaching-learning progress. The difficulty in giving and receiving feedback therefore raises the question: How do we create a harmonious environment between whoever provides the feedback and whoever receives it, in order to make this process easy for the professor and effective for the student?

Typically, upon conclusion of an assessment, students discuss exam questions with their colleagues, and reflect on what they did correctly and incorrectly. Thus the question becomes, “Could personal relationships benefit the process?” and how
can regular feedback among the students in the course be promoted. Learning from colleagues is a practice also known as Peer Instruction (PI), which has brought about important changes in the quality of teaching and has been applied in various universities throughout the world\textsuperscript{1-12}. This active teaching methodology has been found to improve students’ performance and their capacity to resolve problems when they perform activities with their study colleagues.

PI is an interactive teaching-learning process between colleagues, encouraging the student to actively reflect and discuss concepts rather than just receive them passively. If the student were encouraged to think and discuss, these concepts would be more easily assimilated and not learned in a superficial and passive manner to be forgotten over the course of time. In PI, classroom time is used as for brief oral exposition by the professor, followed by a conceptual question, generally presented in a multiple-choice format. The student is instructed to answer the question individually and formulate an argument that justifies his/her reasoning and this answer is captured by a response system that allows the professor to identify the level of student comprehension. After this, the student discusses the question with his/her group of colleagues, learning how to identify and reach agreement on the correct answer. Students who choose different answers to the same question preferably form these groups. The post-discussion answer is also captured by the response system and finally it behooves the professor to make the following decisions: elucidate the question; present a new question; resume the oral exposition of concepts and/or proceed to the next concept to be taught\textsuperscript{57}.

The conceptual questions must not be based on concepts that can be memorized. It is important for the professor to draw up challenging questions, related to the applicability of the concept in practice, and directed towards the objective of the course. The quality of these questions is fundamental to the success of PI. In the classroom, these dynamics require study preparation by the student because, according to Mazur, in the discussion of concepts with colleagues, the first step in the interactive process is to make students read, think and reflect about the concepts before the lesson\textsuperscript{5}.

The study of Crouch and Mazur\textsuperscript{5} presenting ten years of experience with PI, showed that the use of this methodology resulted in a greater capacity for students to resolve problems, when compared with the traditional teaching method, and the studies of Butchart et al.\textsuperscript{5} have suggested that PI may also promote significant improvement in the students’ critical thinking skills. In addition to these student benefits, PI also gives the professor immediate feedback about the level of student comprehension. Thus, the professor is better able to initiate the teaching-learning process, making the necessary adjustments to the rhythm and manner of approach to the concepts.

Therefore, considering that academic exams are generally constructed in accordance with the conceptual questions foreseen in PI and the need for implementing a motivating feedback for both parties in the teaching-learning process, the aim of our study is to identify whether the use of PI on assessment feedback improves the retention of basic concepts in medical programs. There are no systematic studies about the use of PI in assessment feedback.

METHOD

For this study 226 undergraduate students (Y2 = 115, Y3 = 111) enrolled in a Brazilian medical school were invited to participate as a member of either the study or control group.

For each year, we randomly distributed the students into 10 groups with 11-12 students in each group. Based on the students’ general academic performance, we achieved heterogeneity among the groups through random redistribution of members, until all the groups had similar mean academic performance values (overall academic performance between 7.91 and 7.96) and similar numbers of members according to gender. We defined 4 control groups and 4 study groups, interspersed among the topics approached in the course, so that all the students could participate as part of both the control and then the study group.

We selected one second year course (Pharmacology) and two third year courses (Physiology and Anatomy Applied) where these students were enrolled. In each course we randomly selected topics with concepts that are prerequisites for practical activities, namely, Y2: Autonomous Nervous and Cardiovascular Systems; Y3: Skin (lesions, diseases, tumors), Osteoarticular (fractures, bone diseases, anatomic-radiologic correlation) and Anatomy Applied (thorax, spine, pelvis, hip, head and neck). These topics were taught by means of lectures and practical laboratory activities where the student is assessed using conceptual tests with multiple-choice questions.

After taking the regular exam (RE), the students of the control group could receive individual feedback (review of the exam) from the professor according to the course routine, and the students in the study group were invited to participate in the feedback activity. Only during the feedback were the students informed about which teaching method we used.

The study group students had an immediate intervention after the RE with a feedback developed session using the peer instruction teaching method. The feedback groups were formed and a copy of a previously administered exam was given to each student. The students were invited to discuss
the exam questions with their group members. We explained that the goal was for the students to understand the issues regarding their mistakes and correct answers, and thus to better reflect on the contents, assuming that all students had studied and prepared for the exam. The room had computers with Internet access, and students were asked to use them to investigate in greater depth and detail the nature of their mistakes following group discussion with their colleagues. The professor remained in the room the whole time interacting with students when prompted. The time allotted for this activity was unlimited. At the conclusion of the feedback session, the students again answered the post-feedback exam (PFE) so that we could identify any changes in the answers compared with the regular exam taken before feedback. The PFE questions were the same as those in the RE.

Prior to the start of the next academic year (6 months later), we applied a diagnostic exam (DE) to identify whether the students had retained the concepts covered in the previous exams.

The attendance in the feedback between attendees (those who attended the Feedback Session) and absentees (those who had not attended the Feedback Session) was compared in terms of gender, age, performance in the regular exam and overall academic performance of these students throughout the medical program up to that point. We analyzed the performance of students in the RE, PFE and DE and we compared the results between the control group and the study group.

After exploratory and descriptive analyses, the data were analyzed by the Pearson’s chi-square test, independent-samples t-test and linear regression. In all the analyses, the level of statistical significance was set at 0.05%.

## RESULTS

Out of 226 students invited, 101 (44.7%) comprised the study group in our feedback study. Of these students, forty-four were in their second year and fifty-seven in their third year. The remaining 125 (55.3%) students participated as a control group.

Table 1 shows sample characteristics comparing feedback attendees and absentees in terms of gender, age, regular exam performance and overall academic performance (α = .05). There was neither any statistically significant association between gender and feedback attendance nor any statistically reliable difference in terms of age. The scores of regular exam performance were significantly higher for those who had participated in the feedback session than for those who did not and the t-test revealed that the students who participated in the feedback had better overall academic performance compared with the absentees.

Linear regression established that the student’s performance in the RE could significantly predict the performance in the DE, F (2.12) = 27.37, p < 0.0001 and the performance in the RE accounted for 28.8% of the explained variability in the performance in the DE (Figure 1). The regression equation was: predicted DE = 0.114 + 0.370 x (RE).

Figure 2 shows the average correct response score in the exams between control and study groups. The two groups are statistically significantly different in the RE (p = 0.0014) and DE (p < 0.000). The study group demonstrated better performance in both exams than the control group.

We employed the normalized change (c) relation (Marx, 2005) to characterize the students’ performance from RE to PFE. Only two students demonstrated worse performance in the PFE than in the RE and seven had the same performance in both exams. We excluded five students who scored 100% per-

| Table 1 | Samples characteristics comparing feedback attendance in terms of gender, age, exam performance and overall academic performance |
|---------|------------------------------------------------------------------------------------------------------------------------------------------|
|         | Attendance to feedback                                                                                                                                 |
|         | Presents | Absents | t     | df | Sig (2-tailed) | Mean Difference | Lower | Upper |
| Gender (female-male) | n=57, Mean ± DP=22.27 ± 3.52 | n=44, Mean ± DP=22.06 ± 2.03 | -0.544 | 224 | p = 0.587* | -0.20397 | -0.94218 | 0.53425 |
| Age | n=101, Mean ± DP=7.73 ± 1.23 | n=125, Mean ± DP=7.29 ± 1.58 | -2.307 | 224 | p = 0.022* | -0.4427 | -0.82375 | -0.06479 |
| Overall Academic Performance | n=101, Mean ± DP=8.08 ± 0.50 | n=125, Mean ± DP=7.84 ± 0.56 | -3.264 | 224 | p = 0.001* | -0.23423 | -0.37563 | -0.09283 |

*Pearson’s chi-square test; # t-test for independent sample.
We also analyzed each response in the RE and compared them with the DE as shown in Figure 3. Responses that were right in both exams were more frequently found in the study group. In this case there was a huge statistically significant difference (p < 0.000) between the control (mean score = 0.27, SD = 0.18) and study groups (mean score = 0.41, SD = 0.16). Note that a linear regression also established that the study group had more right-right compared to the control group, F (2,12) = 67.94, (p = 0.004).

The study group changed answers slightly more from wrong to right (p = 0.099) than the control group: Study group (mean score = 0.10, SD = 0.09) versus control group (mean score = 0.07, SD = 0.1).

Figure 3 shows the number of answers in the regular exam vs. diagnostic exam for questions that are answered correctly in both exams (right-right), switched from wrong to right, from right to wrong, from wrong to wrong same and wrong to wrong different (study and control group).

The study group changed answers slightly more from wrong to right (p = 0.099) than the control group: Study group (mean score = 0.10, SD = 0.09) versus control group (mean score = 0.07, SD = 0.1).

Figure 3 shows the number of answers in the regular exam vs. diagnostic exam for questions that are answered correctly in both exams (right-right), switched from wrong to right, from right to wrong, from wrong to wrong same and wrong to wrong different (study and control group).

The study group changed answers slightly more from wrong to right (p = 0.099) than the control group: Study group (mean score = 0.10, SD = 0.09) versus control group (mean score = 0.07, SD = 0.1).
As regards the responses that were wrong in both exams, we analyzed separately wrong-wrong different and wrong-wrong same (see Figure 3). Responses that were wrong in RE and wrong different in the DE have no statistically significant difference ($p = 0.819$) but there is a huge statistically significant difference ($p < 0.001$) between study group (mean score $= 0.04$, SD $= 0.03$) and control group (mean score $= 0.18$, SD $= 0.12$) in the responses that were wrong in the RE and after wrong same in the DE. A linear regression model with standardized variables, regardless of RE performance, showed too that the study group had less wrong-wrong same compared to the control group, $F (2,12) = 62.78$, $(p < 0.000)$.

The study groups were potentially much higher in right-right answers and worse in wrong-wrong same answers. The students that had the peer instruction feedback were much less likely to assimilate the incorrect answers than the group that did not attend the feedback session.

None of the students in the control group looked to the professor to receive individual feedback (review of the exam) until the beginning of the next semester.

DISCUSSION

There are many teaching methods and assessment tools utilized in medical schools. Some are inherited from professors that preceded the current faculty or even from their own professors, some are based on experiences reported in the scientific literature, yet others are devised through suppositions regarding what might work. The fact is, although the medical field possesses extraordinary dynamism, lack of investment and a dearth of research in the area has led to criticism that medical education is not in a good state of health.

The Medicine program where this study was developed has a curricular structure composed of integrated courses for six years where, from the 4th year, the student develops practical activities at the primary, secondary and tertiary levels. According to frequent reports by professors, many students begin practical activities without having retained the basic knowledge learned in the previous three years and which are fundamental for the practical activities.

Knowledge of the basic sciences is essential for the practice of Medicine and the retention of this acquired knowledge has been a concern throughout medical education history. Custers and Ten Cate tested for retention of basic science knowledge in medical students and doctors in the Netherlands and their findings do not confirm that most basic science knowledge learned is forgotten quickly. The importance of improving knowledge retention depends on how this knowledge is used or repeated over time and, especially, on whether it is perceived as being essential for the life of the doctor.

Basic science knowledge is lost during the clinical years of medical studies according to Lazic et al. Their study sample included medical students from the second and fifth years with the aim of exploring the level of basic knowledge of physiology and biochemistry, and indicated that clinical knowledge is not based on the knowledge of basic processes. The study by D’Éon et al. also showed that there was considerable knowledge loss among medical students in three basic science courses (Immunology, Physiology, and Neuroanatomy) when he recruited 20 students to retake questions from the first three years compared with their scores 10 or 11 months later.

Custers conducted a review study on long-term retention of basic science knowledge using the Ebbinghaus study published in 1966, which aimed to discover the retention rate of nonsense syllables after different intervals of time. According to the Ebbinghaus curve, after 31 days the retention was 20%.

In our findings, the retention of basic science knowledge (Pharmacology, Physiology and Applied Anatomy), after six months was 24% of the received knowledge. If we consider that this knowledge is important to the development of medical practice in the next steps of the course, these results are lower than we had expected. These data suggest that this knowledge was nonsense for our students or that they failed to perceive its importance to their medical practice.
to predict that the immediate feedback after the evaluation of a skill is more effective. Phye and Andre found that students who received immediate feedback repeated fewer errors in the post-test than students who received the delayed feedback. It is important to create mechanisms so that the student does not assimilate imprecise concepts. The lack of feedback can lead the students to interpret improperly their learning process and may develop a false confidence and a repetition of misconceptions.

In our study, when we gave feedback, using PI immediately after the exam, retention of basic science knowledge jumped to 39%, increasing by 15%. The students that had assessment feedback had the opportunity to discuss their misconceptions. These students had the highest number of correct answers with assimilated knowledge and fewer assimilation of wrong answers, therefore, students who received immediate feedback had less tendency to make the same conceptual errors.

Mazotti et al., in a study on the perception of longitudinal versus traditional assessment in medical internship, found that between the students and tutors there was greater availability to provide and receive feedback, suggesting that this finding is related to the involvement between the two parties. The authors suggested that this relationship might favor the emotional side of the process. As in the studies of Bates et al., the students’ relationship with the working team was positive after providing the tutors with feedback. The students described a feeling of support, affection and security. In our study, we used the Peer Instruction method where the students provided and received feedback from their peers, their colleagues.

Rao and DiCarlo used PI in their studies and concluded that students demonstrate a significantly better performance in multiple-choice questions following discussion with their colleagues. The authors suggested that this improvement might also be explained by the studies of Silberman, who stated that when a person is capable of teaching another, it is because he/she has mastered the concept.

PI in feedback was found to be effective in improving the retention of basic science knowledge. This method of active learning and collaborative teaching promotes better performance among students. The relations between the students during the discussion motivated the working team, essential to the understanding and retention of medical formation.

REFERENCES
1. Cavalcani RB, Detsky AS. The education and training of future physicians: why coaches can’t be judges. JAMA 2011; 306(9): 993-994.
2. Norcini J. The power of feedback. Medical Education 2010; 44: 16-17.
3. Ridder JMM, Stokking KM, McGaghie WC, Cate OT J. What is feedback in clinical education? Medical Education 2008; 42:189-197.
4. Bienstock JL, Katz NT, Cox SM, Hueppchen N, Erickson S, Puscheck EE. To the point: medical education reviews - providing feedback. American Journal of Obstetrics & Gynecology 2007; 508-513.
5. Mazur E. Peer instruction: A user’s manual. Prentice Hall, Upper Saddle River, NJ 1997.
6. Araújo IS, Mazur E. Instrução pelos colegas e ensino sob medida: uma proposta para o engajamento dos alunos no processo de ensino-aprendizagem de física. Cad Bras Ens Fís 2013; 30(2):362-384.
7. Lasry N, Mazur E, Watkins J. Peer instruction: from Harvard to the two-year college. American Journal of Physics 2008;76:1066.
8. Crouch CH, Mazur E. Peer instruction: ten years of experience and results. American Journal of Physics 2001;69(9):970-977.
9. Butchart S, Handfield T, Restall G. Using peer instruction to teach philosophy, logic, and critical thinking. Teaching Philosophy 2009;32:1-40
10. Rao SP, DiCarlo SE. Peer Instruction improves performance on quizzes. Advances in Physiology Education 2000; 24(1):51-55.
11. Silberman M. Active learning: 101 strategies to teach any subject. Boston: Allyn and Bacon, 1996.
12. Fagen AP, Crouch CH, Mazur E. Peer instruction: results from a range of classrooms. The Physics Teacher 2002;40:206-209.
13. Whitehead, C. What ails medical education? Science 2015; 347: 241.
14. Jason H. The relevance of medical education to medical practice. JAMA 1970;212:2092-5.
15. Blizard PJ, Carmody JJ, Holland RAB. Medical students’ retention of knowledge of physics and biochemistry on entry to a course in physiology. Br J Med Educ 1975;9:249-54.
16. Kennedy WB, Kelley PR, Safran M. Use of NBME examinations to assess retention of basic science knowledge. J Med Educ 1981;56:167–73.
17. Sukkar MY. An approach to the question of relevance of medical physiology courses. Med Educ 1984;18:217–21.
18. Conway MA, Cohen G, Stanhope N. On the very longterm retention of knowledge acquired through formal educa-
tion: twelve years of cognitive psychology. J Exp Psychol
Gen 1991;120:395–409.
19. Swanson DB, Case SM, Luecht RM, Dillon GF. Retention of
basic science information by fourth-year medical students.
Acad Med 1996;71 (Suppl):80–2.
20. Custers, EJ, Ten Cate OT. Very long-term retention of ba-
sic science knowledge in doctors after graduation. Med
Educ. 2011;45(4):422-30.
21. Lazic,E, Dujmovic, J & Hren,D. Retention of basicscience
knowledge at clinical years of medical curriculum. Croa-
tian Medical Journal, 2006; 47:882-997.
22. D’Eon,MF. Knowledge loss of medical students on first
year basic science courses at the University of Saskat-
chewan. BMC Medical Education, 2006; 6(5).
23. Custers, EJFM. Long-term retention of basic science know-
ledge: a review study. Adv in Health Sci Educ, 2010; 15:109-
128.
24. Ausubel, DP. The acquisition and retention of knowledge: a
cognitive view. Kluwer Academic Publishers, 2000. P.212.
25. Farquharson AL, Cresswell AC, Beard JD, Chan P. Rando-
mized trial of the effect of video feedback on the acquisi-
tion of surgical skills. British Journal of Surgery 2013; 100:
1448-1453.
26. Thomas JD, Arnold RM. Giving feedback. Journal of Pallia-
tive Medicine 2011; 14(2): 233-239
27. Hodder RV, Rvington RN, Calcutt LE, Hart IR. The effect-
iveness of immediate feedback during the objective struc-
tured clinical examination. Medical Education 1989; 23:184-
188.
28. Phyed, Andre T. Delayed retention effect: attention, per-
severation, or both? Contemporary Educational Psychology
1989;14(2):173-185.
29. Mazotti L, O’Brien B, Tong L, Hauer KE. Perceptions of
evaluation in longitudinal versus traditional clerkships.
Med Educ 2011;45:464–70.
30. Bates J, Konkin J, Suddards C, Dobson S, Pratt D. Student
perceptions of assessment and feedback in longitudinal in-
tegrated clerkships. Medical Education 2013;47:362-374.

ACKNOWLEDGEMENTS
The authors thank: Helen M. Shields, Kelly Miller, Orad Reshef
for comments that greatly improved the study and Eliane M.
Ingrid Amstalden, André Schenka and Lindemberg Silveira
Filho for support with the data.

FINANCIAL SUPPORT
Coordination for the Improvement of Higher Education Per-
sonnel (Capes), Ministry of Education, Brazil.

ETHICAL APPROVAL
Ethical approval for this study was obtained from Ethics Com-
mittee. CAAE: 42445315.7.0000.5404 (May, 28, 2015).

CONTRIBUTIONS
SMRRP: wrote the manuscript, provided data in the study and
takes responsibility for the integrity of the data and the accu-
ragy of the data analysis. EM: has made a substantial contribu-
tions to the conception, design of the work and all the process.

CONFLICT OF INTERESTS
The author(s) declare(s) that there is no conflict of interest re-
garding the publication of this article

POSTAL ADDRESS
Silvia Maria Riceto Ronchim Passeri
Rua Araticum, 39 – Campinas – São Paulo
Cep 13098-368 – spasseri@unicamp.br

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits
unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.