When is team exam performance a better measure of learning than individual perfor...

Carmichael (2009) reported that team-based learning enhances individual performance in an introductory biology course. He also found that teams perform significantly better on quizzes and exams than even the students, on the teams, whose individual performances on the same quizzes and exams were the highest of all team members. Such findings are consistent with a growing literature showing that teams out perform every member of the teams on quizzes and exams (Michaelsen et al., 2004 ; Nieder et al., 2005 ; Vasan et al., 2008 , 2009 ; Chung et al., 2009 ; Wiener et al., 2009). Furthermore, all team members likely contribute to team performances on exams (Van Winkle et al., 2011). Team performances on quizzes and exams are usually about 10% higher on average than scores of the top team members. For these reasons, it is probably fair for all team members to receive the same scores on team exams, as also concluded by Carmichael (2009) .

What should we conclude, however, when team performances greatly exceed those of individual members of the team? We encountered such a team in an introductory biology course at Arapahoe Community College (Colorado) in the fall semester of 2016. The three members of this team averaged 38, 39, and 65% on six quizzes (percent of multiple choice questions answered correctly) when team-based learning was introduced in the course over the second half of the semester. In the team-based learning setting, students complete quizzes/exams as teams, immediately after they finish the same quizzes/exams as individuals. The team highlighted here, however, averaged 91% for the same quizzes on which they had just performed rather poorly as individuals.

Estimates of effect size (ES) are useful ways to determine the magnitude and importance of such differences. One estimate of ES is the Pearson correlation coefficient (r), which can be calculated from the t statistic (Rosenthal and Rubin, 1982 ; Hojat and Xu, 2004). In the present case, a t -test was performed by first calculating the population of differences between the team performances on each of the six quizzes and the best individual performances by a team member. These six differences were then used in a paired t -test to calculate the value of the t statistic (GraphPad Prism 5 Software, Inc., La Jolla, CA, USA), which was highly statistically significant (p = 0. 002). Moreover, the ES, for this better performance by the team than their best individual, was of crucial practical importance (r = 0. 93) (Hojat and Xu, 2004). That is, the ES was statistically equivalent for finding a drug that lowers the death rate from 96. 5 to 3. 5% (Rosenthal and Rubin, 1982). When only three quizzes, that all three members failed (i. e., scores of 56% or lower), are considered, the ES was 0. 98 (p = 0. 02). In the latter case, the average individual performances were 37, 40, and 50%, whereas the team scores averaged 85%.

How does one reconcile team learning in the middle of the " B" range for most colleges in the USA, while team members fail the same quizzes rather badly? Do individual performances reflect failed learning in this case, or did students collectively really learn quite a bit? All three students on this team were struggling to earn a " C" in the course, so it is unlikely that they divided up the work to study for the team quizzes when their individual scores would have surely suffered. Moreover, there was no easy way for them to cheat, and their discussions while answering questions on the team quizzes were

much more highly animated than other teams in the course. In fact, this team always finished their work on team quizzes at least 5 min after all of the other teams completed the quizzes.

In addition, a better representation of learning might be to include partial credit when a team answers a multiple choice question correctly on their second or even third attempt. In this case, the team discussed above averaged 96% on all six quizzes and 92% on the three quizzes that all three members failed as individuals. Furthermore, the overall performance by this team could not be distinguished, on statistical grounds, from another top team whose membership included the only " A" student in the class of 24 students. Opportunities for students to earn partial credit in this way have also been linked to increased learning and retention of course material (Epstein et al., 2002 , 2003 ; Cortright et al., 2003 ; Dihoff et al., 2003 , 2005 ; Cotner et al., 2008).

In the end, the performances of this and similar teams raise questions I find impossible to answer adequately. Our culture in the USA has taught us that if individual students earn 90s on their quizzes and exams, they deserve grades of As for their efforts. Does this dogma also apply to the performances of teams of students? If teams of students earn 90s on their team quizzes and exams, do they also deserve As for their work, even or especially when their individual scores are very low? Are we reluctant to disregard their individual scores and award them As for such efforts, because we cannot submit grades for teams? Or is it because we value measures of individual learning highly, because of our cultural indoctrination? We have no

such indoctrination to guide us when it comes to high team performances, in the face of otherwise poor performances by individual members of the teams.

In this regard, it is worth addressing one additional bias about the team discussed above. I decided half way through the biology course to switch to a team-based learning format in the time designated for lecture, because I was disappointed in students' performances on lecture exams during the first half of the course. I had used team-based learning in previous courses to improve students' performances on quizzes and exams (e. g., Van Winkle et al., 2011), so it seemed to be a quick way to resolve my concern. Since the students were already in teams for the laboratory portion of the course, I decided to keep them in the teams in which they had already developed some rapport. I was worried about one team, however. Members of the team discussed above not only had failing or near failing grades on both lecture exams up to that point in the course but also frequently had disagreements in the lab and complained to me privately about each other. My plan was simple; I would take them aside and ask them if they wanted to remain a team or break up and join other teams. I even had decided on the other teams I would assign them to. To my surprise, they wanted to remain together. Their " A" work together on team lecture quizzes, thus, became even more beautiful to behold.

Beautiful, in part, because little is known empirically about how teams perform so much better on exams than even the best individual performers on teams. For example, Michaelsen et al. (2004) stipulate precisely how to

form successful student teams, but little of this prescription was used in the present case. Similarly, students have been surveyed to learn their beliefs on how team-based learning influenced their learning styles (Oldland et al., 2016), but nothing was learned about how their team exam performances exceeded performances by every member of the teams. Still, we know that not only teams perform better than all individuals on teams but also the poorest performing members of mature teams likely contribute to team exam performances (Van Winkle et al., 2011). So my dilemma remains; is it better to reward application of apparently limited individual knowledge to achieve superior team work (as the team under discussion did very well), or simply to measure apparently limited individual knowledge?

Author Contributions

The author confirms being the sole contributor of this work and approved it for publication.

Conflict of Interest Statement

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Carmichael, J. (2009). Team-based learning enhances performance in introductory biology. *J. Coll. Sci. Teach.* 38, 54–61.

Google Scholar

Chung, E. K., Rhee, J. A., Baik, Y. H., and OS, A. (2009). The effect of team-based learning in medical ethics education. *Med. Teach.* 31, 1013–1017. doi: 10. 3109/01421590802590553

PubMed Abstract | CrossRef Full Text | Google Scholar

Cortright, R. N., Collins, H. L., Rodenbaugh, D. W., and DiCarlo, S. E. (2003). Student retention of course content is improved by collaborative-group testing. *Adv. Physiol. Educ.* 27, 102–108. doi: 10. 1152/advan. 00041. 2002

PubMed Abstract | CrossRef Full Text | Google Scholar

Cotner, S., Baepler, P., and Kellerman, A. (2008). Scratch this! The IFAT as a technique for stimulating group discussion and exposing misconceptions. *J. Coll. Sci. Teach.* 37, 48–53.

Google Scholar

Dihoff, R. E., Brosvic, G. M., and Epstein, M. L. (2003). The role of feedback during academic testing: the delay retention effect revisited. *Psychol. Rec.* 53, 533–548.

Google Scholar

Dihoff, R. E., Brosvic, G. M., Epstein, M. L., and Cook, M. J. (2005). Adjunctive role for immediate feedback in the acquisition and retention of mathematical fact series by elementary school students classified with mild mental retardation. *Psychol. Rec.* 55, 39–66.

Google Scholar

Epstein, M. L., Brosvic, G. M., Costner, K. L., Dihoff, R. E., and Lazarus, A. D. (2003). Effectiveness of feedback during the testing of preschool children, elementary school children, and adolescents with developmental delays. *Psychol. Rec.* 53, 177–195.

Google Scholar

Epstein, M. L., Lazarus, A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., et al. (2002). Immediate Feedback Assessment Technique promotes learning and corrects inaccurate first responses. *Psychol. Rec.* 52, 187–201.

Google Scholar

Hojat, M., and Xu, G. (2004). A visitor's guide to effect size: statistical versus practical significance of research findings. *Adv. Health Sci. Educ.* 9, 241–249. doi: 10. 1023/B: AHSE. 0000038173. 00909. f6

CrossRef Full Text | Google Scholar

Michaelsen, L. K., Knight, A. B., and Fink, L. D. (2004). *Team-Based Learning: A Transformative Use of Small Groups in College Teaching* . Sterling, VA: Stylus Publishing.

Google Scholar

Nieder, G. L., Parmelee, D. X., Stolfi, A., and Hudes, P. D. (2005). Team-based learning in a medical gross anatomy and embryology course. *Clin. Anat.* 18, 56–63. doi: 10. 1002/ca. 20040

PubMed Abstract | CrossRef Full Text | Google Scholar

Oldland, E., Allen, J., and Currey, J. (2016). Students' perception of the role of team-based learning in shaping individual learning style, team skills and clinical practice. *Aust. Crit. Care* 29, 117. doi: 10. 1016/j. aucc. 2015. 12. 019

CrossRef Full Text | Google Scholar

Rosenthal, R., and Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *J. Educ. Psychol.* 74, 166–169. doi: 10. 1037/0022-0663. 74. 2. 166

CrossRef Full Text | Google Scholar

Van Winkle, L. J., Chandar, N., Green, J. M., Lynch, S. M., Viselli, S. M., and Burdick, P. (2011). Does critical reflection by biochemistry learning teams foster patient-centered beliefs among medical students? *Med. Sci. Educ.* 21, 158–168. doi: 10. 1007/BF03341613

CrossRef Full Text | Google Scholar

Vasan, N. S., DeFouw, D. O., and Compton, S. (2009). A survey of student perceptions of team-based learning in anatomy curriculum: favorable views unrelated to grades. *Anat. Sci. Educ.* 2, 150–155. doi: 10. 1002/ase. 91

PubMed Abstract | CrossRef Full Text | Google Scholar

Vasan, N. S., DeFouw, D. O., and Holland, B. K. (2008). Modified use of team-based learning for effective delivery of medical gross anatomy and embryology. *Anat. Sci. Educ.* 1, 3–9. doi: 10. 1002/ase. 5

PubMed Abstract | CrossRef Full Text | Google Scholar

Wiener, H., Plass, H., and Marz, R. (2009). Team-based learning in intensive course format for first-year medical students. *Croat. Med. J.* 50, 69–76. doi: 10. 3325/cmj. 2009. 50. 69

PubMed Abstract | CrossRef Full Text | Google Scholar