Assessing the Effect of Test Speededness on Individual and Collaborative Exams

Owen H. Hudson[1,a], Noah Benscher[1,2,a], Naomi Greenberg[1], Malia M. Fortney[1], Morgan Byron[1], Rosalie Koenig[2], Alexandra Bitton-Bailey[3], Samuel J. Martins[1]*

[1]Department of Plant Pathology, University of Florida, Gainesville, FL, 32611, USA
[2]Agronomy and Center for Stress Resilient Agriculture, University of Florida, Gainesville, FL 32611, USA.
[3] Center for Teaching Excellence, University of Florida, Gainesville, FL, 32611, USA

[a] These authors contributed equally to this study
* Corresponding author: sj.martins@ufl.edu

Abstract:
Identifying strategies that improve learner's information retention, collaboration skills, and academic performance are of acute interest in higher learning. Additionally, modeling real-world scenarios to prepare students for post-academic life can be greatly beneficial. By using different testing strategies, such as assigning exams to be done individually or in groups and modulating the allotted time-per-question, one can compare the efficacy of these changes to the standard testing format. In this study, we used data from two undergraduate courses to evaluate how changing the time per question and how working in groups impacts test grades. Self-reflection questions were also used to gauge student's satisfaction with the changes and see if their perceptions aligned with the data. On average, students' performance improved when taking tests in groups compared to individually, however, the increase in time-per-question did not have a consistent impact. Students had overall positive reactions to working in groups with a greater sense of confidence and collaboration reported in the self-reflection questionnaire. While more data are required to make definitive conclusions, these data show that integrating additional group-based activities, specifically test-taking, can improve student performance and enjoyment in learning outcomes.

Keywords: Collaborative learning, active learning, trivia, games, test speededness

Introduction

Collaborative learning (CL) is an educational approach by which learners work with others (students, peers, teachers) to problem solve and understand concepts. This method contrasts with traditional, teacher-centered instruction by including students in a more active learning process. The positive effects of CL on students' development and performance are supported by decades of research. Implementing CL practices has been shown to greatly improve student individual learning (Linton et al., 2014), engagement (Molinillo et al., 2018), and satisfaction (Le et al., 2017; Martins et al. 2021) in the classroom. While CL includes a diverse array of specific lesson plans and methods, it can be used in all subjects and is relevant to all aspects of life. In academia, students work together when studying and completing coursework. Collaboration persists outside academia in all careers as people communicate and work together to fulfill individual and collective goals. Collaboration also facilitates work efficiency and time management to meet deadlines. Certain CL techniques and activities mimic real-world scenarios and therefore can prepare students beyond the teacher-centered alternative (Cheng et al. 2021).

Test speededness is the measure of how time constraints on tests affect the performance of the individual taking the test (Cintron 2021). Different testing scenarios highlight the complex nature of test speededness and the difficulty in generalizing across groups, but also why it could be a valuable metric to monitor and explore. For example, compare a test that consists of completing as many basic arithmetic questions as a test-taker can in 30 seconds to a test where the test-takers have several hours to complete it, but the complexity of each question varies greatly. These two scenarios provide different information about learning outcomes and the individuals taking the test. There is limited research on the impact of CL on test speededness. Our preliminary research found that collaborative testing improves student performance and speed on an exam (Greenberg & Martins 2022) but did not include exams with a time constraint.

Collaborative testing is a concept that connects CL and test speededness, allowing students to work with others on an exam. To further explore collaborative testing, this study aimed to evaluate the effect of collaborative testing compared to individual testing on test speededness. We present data on students' perceptions of self-efficacy in these two testing settings and the extent to which students learned something new while taking a test in a group. We also used the virtual, game-based learning platform "Kahoot!" to perform a similar evaluation on test-takers when given varying levels of per-question time. Results from previous classes found that while the time taken to complete an exam individually did not predict the student's grade, a negative correlation ($P=0.0416$ $r=-0.351$) was found between student's grades and the time it took to complete the exam as a group (Supplemental Figure 1). Groups that completed the exam faster tended to have higher grades, suggesting that students work more efficiently and achieve higher results when collaborating. Additionally, it was found that students perform better individually in Kahoot! games when there is a 20 second time constraint per question compared to a 30 second time constraint (Supplemental Figure 2). Based on these previous results, we hypothesized that group learning enhances student performance and efficiency. The objectives of this study were to:

1) Measure the effect of collaborative versus individual test-taking.
2) Measure the effect of test speededness on accuracy.
3) Measure the effect change of test speededness on collaborative versus individual test taking.
4) Analyze how students' self-perception of performance reflects their actual performance.

Materials and Methods

Class Design

The study consisted of data from two General Education undergraduate courses at the University of Florida: “PLP2000: Plants, Plagues, and People” and “PLP2311: What Are Plants Talking About?” that were taught simultaneously by the same instructor in the Spring of 2023. One class had 14 students, and the other class had 43 students. The same instructor taught both courses.

Two-Stage Exams

Two two-stage exams were completed during the semester and were divided into individual and group segments. Students had a three-hour block to complete both sections, and because no student exceeded two hours, we considered this the non-time-constrained exam. Initially, students completed the exam individually. Once all students had finished the individual portion and handed it in, students formed groups of four or five members on their own to retake the same exam collaboratively. Collaborative work was done by communicating quietly in groups on a single exam copy. The groups were distantly arranged in the classroom and instructor and TA walked around the room to make sure groups were not trying to listen to the discussion from other groups. The groups varied from week to week. The two-stage exams consisted of 20 questions. 10 questions were created by the course instructor and the other 10 questions were selected by the instructor from the questions created by students for the weekly quizzes. Student exam grades were 60% attributed to individual performance, 30% attributed to group performance, and 10% attributed to completion of the post-exam self-reflection questionnaire.

Kahoot! Quizzes

Kahoot! provided a virtual, game-based testing platform that students could log into and answer from their phones or computers. Kahoot! quizzes were administered as time-constrained exercises that were delivered in two parts, individually and in a group. The total number of questions per exercise was 20, consisting of 17 multiple choice questions and 3 true/false questions. The questions were different for the individual and group scenarios (two sets of 20 questions), and a time constraint was imposed on each question. One quiz had a time constraint of 30 seconds per question for individual and group parts. The second had a time constraint of 20 seconds per question for individual and group parts. If the student or group did not choose an answer for a question, the question would be marked as incorrect. The score is formulated in Kahoot! by the quickness and accuracy of students’ answers, with higher scores achieved for correct answers given more quickly. An incorrect answer yields a score of 0. The correct answer was immediately provided to the students after each question. Groups consisted of 4 - 5 students. For all Kahoot! exercises, the individual part was administered before the group part. Student performance on the Kahoot! exercises did not affect students’ grades in either class. There were four total iterations of the Kahoot! quizzes, two for each class, resulting in eight separate evaluations. Kahoot! exercises were administered during the class period prior to the class of the two non-time-constrained exams.

Self-Reflection Questionnaire

Self-reflection questionnaires were administered immediately after students completed a two-stage exam. These surveys consisted of the following 14 questions (Table 1) and were returned anonymously. Students received 10% of their grades by simply returning the questionnaire with an answer. Students were informed before completing the questionnaire that there were no right or wrong answers. We collected qualitative data from the surveys as well as quantitative data from numeric-scale questions.

Table 1: List of questions used in student self-reflection questionnaires.

Question Number	Question
1	Describe your experiences and thoughts about completing the test individually as compared to in a group.
2	Did you learn anything from your colleagues during the collaborative part of the test that was not clear before, or were you already comfortable with the subject?
3	Did you find it faster to do the test in a group or by yourself? If yes, why? If no, why?
4	Did you find it easy to collaborate with your peers? If yes, why? If no, why?
5	Did you experience any challenges when completing the test with your group? Please explain.
6	Did any of the questions you created show up on the test? If yes, did you remember the answer?
7	Explain why you agree or disagree with the following statement: Creating questions helped me to learn the subject.
8	On a scale of 1 to 10, how well do you perform when answering timed questions? 1 = very bad; 10 = great.
9	On a scale of 1 to 10, how competitive do you consider yourself to win a game (example: trivia)?
10	What challenges (if any) do you perceive to have the greatest influence on your performance on timed questions?
11	On a scale of 1 to 10, how closely does your perception match your performance (about question 10)? 1 = never; 10 = always
12	Did you like doing Kahoot! before the exam? If yes, why? If no, why?
13	On a scale of 1 to 10, how do you rate the difficulty of this exam? 1 = too easy; 5 = fair level; 10 = too hard
14	What is your gender identity? (select all that apply): Woman, Man, Transgender, Non-binary/non-conforming, Prefer Not to Answer

Data Collection and Data Analysis

All activities conducted in this research were approved by the Institutional Review Board, with the service survey number IRB202100054. Wilcoxon signed-rank test (Whitley & Ball 2002) was used to assess the Kahoot! and exam performances. Data analysis was performed in RStudio and SigmaPlot® version 14.5, RStudio (RStudio Team, 2020), and wordcloud package (Ian Fellows, 2018) were used for data analysis and artwork.

Results

Time-Constrained Evaluations

Kahoot! quizzes served as time-constrained evaluations. Group performance was better than individual performance on all Kahoot! evaluations ($P < 0.001$ for 20s/question; $P = 0.004$ for 30s/question). The average individual performance on the 20 second Kahoot! quizzes were 67% in one class and 61% in the other class. The average group performance for these classes was 73% and 80%, respectively. For Kahoot! quizzes with the 30 second time constraint, average individual scores were 60% and 55% and average group scores were 78% and 85%, respective of each class. Without grouping by class, the average individual performance on the Kahoot! evaluations with a time constraint of 20 seconds was 66.4% and average group performance was 75.6%. Under a 30 second time constraint, average individual performance fell to 58.5% but average group performance increased to 79%.

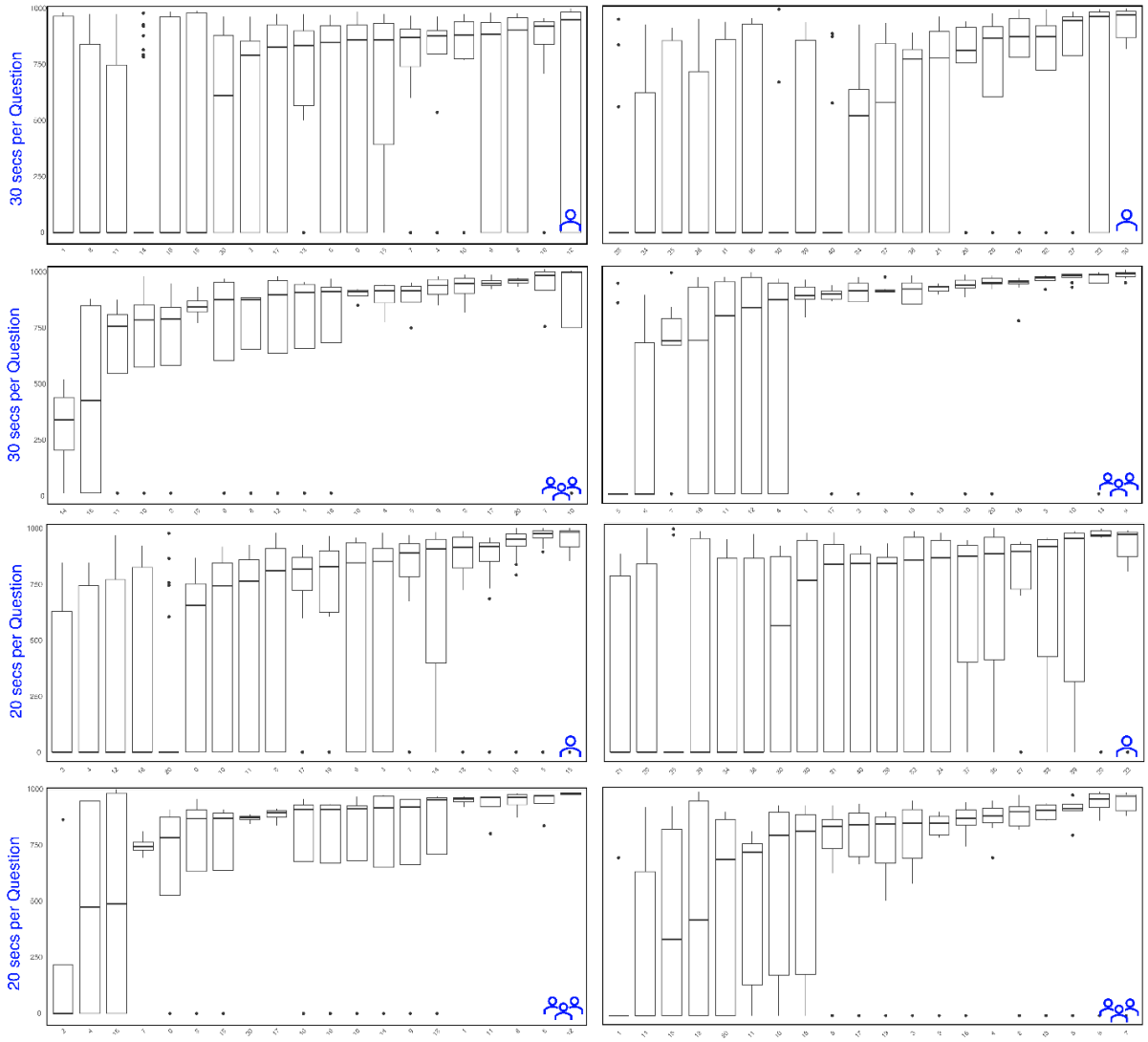


Figure 1: Individual and group performance on Kahoot! evaluations. Individual performance and group performance are indicated by the silhouettes in the bottom right of each box plot. The x axis of each graph represents question number, and the y axis is the possible score. Question number is arranged by increasing median score. Repetitions of Kahoot! evaluations by different student populations are shown in adjacent box plots. Kahoot! evaluations were administered with a time constraint of 20 seconds per question and 30 seconds per question. Individuals and groups had completely different sets of questions, all related to the same subject. Results from the top 4 graphs were completed by students from the PLP2311 course while the bottom 4 graphs were completed by students from the PLP2000 course.

On average, groups took less time than individuals did to answer a question correctly for all four Kahoot! evaluations (Figure 1). This trend (groups taking less time than individuals) was present for all iterations of the Kahoot! evaluations, regardless of whether there was a 20-second or 30-second time constraint per question.

Two-Stage Exams

When individual performance was compared to group performance for the exams, better average results were obtained for the group scores than individual scores on all the two-stage exams ($p < 2.74e-06$, $p < 9.01e-07$, $p = 0.006$, $p = 0.001$) (Figure 2). Specifically, 75% of students performed better on exams when working in groups than working individually. No student performed worse while working in a group on exam 4.

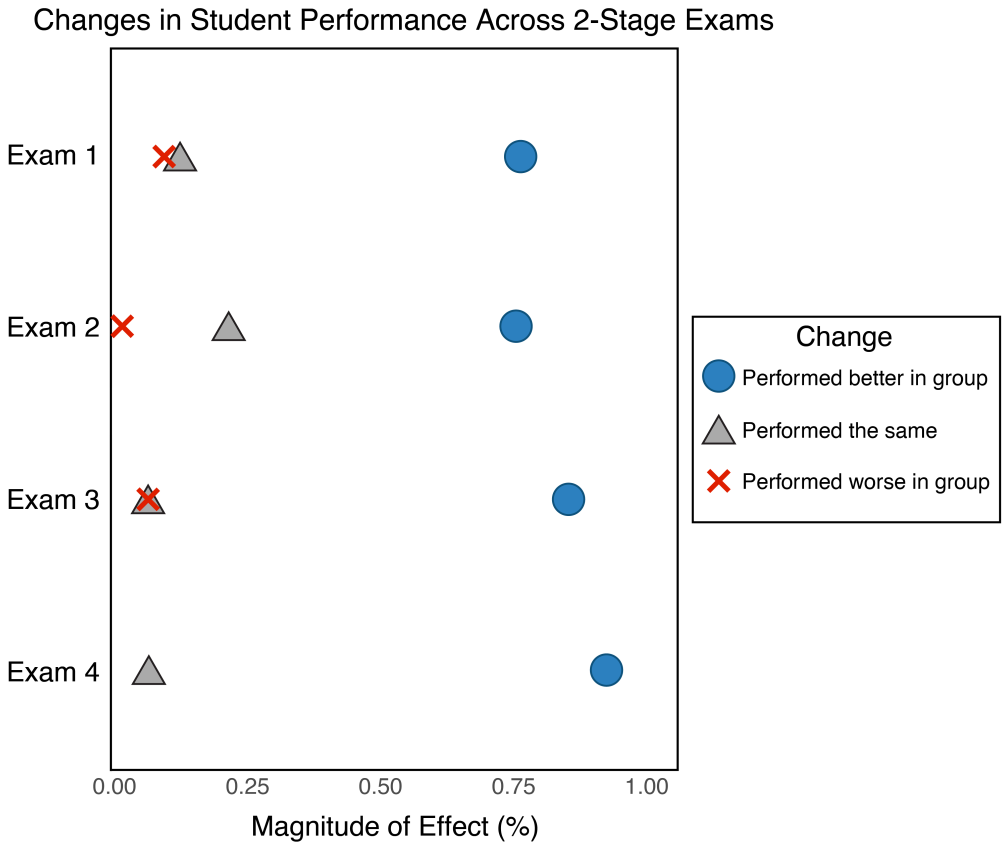


Figure 2: The effect of exam setting (individual vs group) on student performance was measured using a Wilcoxon signed-rank test. In this figure, performance is categorized by "performed better in group," "performed the same," and "performed worse in group." Exams 1 and 2 were completed by students from the PLP2311 course while exams 3 and 4 were completed by students from the PLP2000 course. The proportion of students that performed better in a group setting was the highest on all four exams with a magnitude of ≥ 0.75. No student performed worse in a group on Exam 4, so there is no red X present.

All exams, except for exam 3, saw a reduction in the average time taken to complete the exam. This can be seen by comparing vertical dashed lines across corresponding plots in Figure 3. Even though exam 3 took longer in groups than individually, it was only by one minute. Individual and group performance can be compared by looking at the horizontal dashed lines across corresponding plots which indicates average performance.

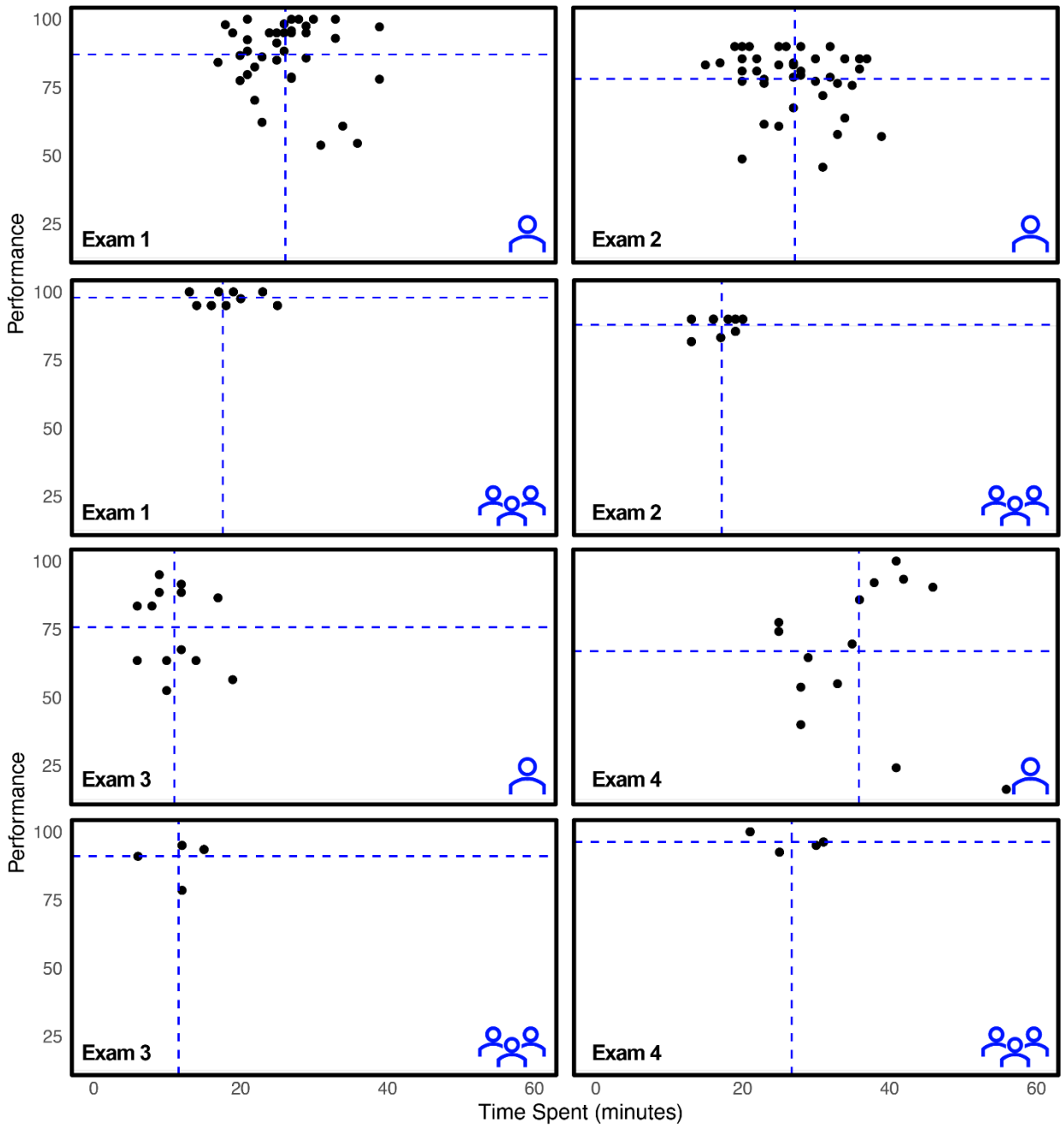


Figure 3. Performance and time to complete the exam individually and in groups for 4 exams in two general education courses. Exams 1 and 2 were completed individually (single person icon) or in groups (3 people icon) by students from the PLP2311 course while exams 3 and 4 were completed individually (single person icon) or in groups (3 people icon) by students from the PLP2000 course. Horizontal dashed lines indicate average performance per exam, and vertical dashed lines indicate average time taken per exam.

Self-Reflection Questionnaire

Students had mostly positive remarks regarding taking tests as a group compared to doing so individually as reported in all the self-reflection questionnaires (two questionnaires per class): 74% of responses to question 1 preferred working in groups with the most common reasons being

reduction in time and stress and the ability to collaborate with others. Twenty one percent (21%) had no preference taking the exams individually or in groups, and 5% preferred taking the exams individually. Similarly, only 3% of students responded to question 3 that they felt working in groups was slower, and only 1% responded to question 4 that they found working in groups was not easier (Figure 4). Students found that group testing allowed them to share and discuss ideas which reportedly improved their confidence. For instance, some students reported: “I felt more confident taking the test as a group and felt like I learned more from my peers”. Others found group testing to be less stressful and helped reduce their anxiety: “The test was very stressful individually, so when I worked in a group I felt a lot more calm”. Furthermore, over 82% of survey responses to question 2 indicated that students either learned something new or received clarification on material that they were unsure about.

Figure 4. Wordclouds for responses to questions 1, 3, and 4 of the self-reflection questionnaires.

We investigated why students performed better in groups than individually on the self-reflection questionnaires, specifically questions 8-12 (Table 1). There was no significant correlation between the change in group and individual performance for each student (dependent variable) and questions 8, 9, and 11 (independent variables) (P= 0.434, 0.170, 0.231), suggesting that additional evidence is required for the quantitative survey questions.

The most common challenges identified in question 10 were time pressure (43%), lack of confidence (21%), anxiety (18%), and lack of knowledge (13%).

Fifteen percent (15%) of the students reported facing challenges when working in a group (Question 5). Those who did report challenges described the difficulty arising when other students

were confident about certain answers that they disagreed with. This led to peer-to-peer discussion as all groups finished and submitted the exam without instructor intervention.

Discussion

The results of this study highlight the benefits of group testing over individual testing in both time-constrained and non-time-constrained evaluations. The data showed that group performance was consistently better than individual performance in both Kahoot! quizzes and two-stage exams. Specifically, the average group performance was greater than the average individual performance regardless of the time constraints. The results indicate that groups not only achieved higher scores but also took less time to answer correctly, suggesting greater efficiency and accuracy in a collaborative setting. Despite the small sample size of 57 students from only two classes, the significant trends observed in this study warrant further investigation. The greater efficiency is likely due to the collaborative nature of group testing, which facilitates idea sharing and collective problem-solving.

The self-reflection questionnaire provided further insights into the perceived benefits of group testing. Students reported that group testing allowed them to share and discuss ideas, which improved their confidence and reduced anxiety: “I enjoyed taking [the exam] as a group because my team had several ideas I didn’t consider”. Over 82% of the survey responses indicated that students learned something new or received clarification on material they were unsure about, showing the potential value of peer discussions during group tests. This finding aligns with previous research that has shown students benefit from peer instruction, which enhances their understanding and retention of course material (Tullis & Goldstone, 2020). However, it is essential to refine the survey questions to better understand the challenges students face when working in groups. Previous studies have shown that test anxiety does not diminish when taking tests as a group compared to individually (Breedlove et al., 2004). Yet, the structure of individual and collaborative testing in this study differs from those in previous research, indicating a need for more tailored studies.

Interestingly, the data also revealed that individuals performed better under stricter time constraints, suggesting that the pressure of limited time might enhance focus and performance. This observation was consistent across both the preliminary and presented data. However, it is worth noting that not all students participated equally in the Kahoot! evaluations. The different participation levels indicate that either (or both) time constraints or lack of impact on class grade might have diminished engagement. Less pressure on any assignment or exam because the result will not impact their overall grade likely leads to less effort or focus during those ungraded assignments/exams.

The challenges identified in the self-reflection questionnaire, such as time pressure, lack of confidence, and anxiety, were common among students. These issues suggest that additional research is needed to explore how different time constraints impact student performance and stress levels. Manipulating the time constraint from 20 and 30 seconds to 20 and 60 seconds per question could provide valuable insights into alleviating potential time-pressure experienced by students. The positive reception of group testing by students, with 100% of survey responses affirming the value of Kahoot! evaluations for reviewing course material and preparing for exams, underscores the potential benefits of incorporating collaborative testing methods in educational

settings. This study's findings echo the results of existing literature that associates a higher sense of belonging in science classrooms with higher academic achievement and persistence (Walton & Cohen, 2011; Yeager & Walton, 2011). Additionally, the extent to which students contribute to discussions plays a crucial role in their learning experience (Penuel et al., 2023).

In conclusion, this study provides compelling evidence that group testing can enhance student performance, confidence, and learning. However, further research with larger sample sizes and varied time constraints is necessary to validate these findings and address the identified challenges. By refining the research design and survey instruments, future studies can offer deeper insights into the dynamics of group versus individual testing in educational contexts.

Data availability statement

The dataset is freely available after registration as a researcher. Computing workflows (source code and input files) are available on GitHub (https://github.com/Martins-Lab/WordCloud).

Conflict of interest

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

Cheng, F. F., Wu, C. S., & Su, P. C. (2021). The impact of collaborative learning and personality on satisfaction in innovative teaching context. *Frontiers in Psychology*, *12*, 713497.

Cintron, D. W. (2021). Methods for Measuring Speededness: Chronology, Classification, and Ensuing Research and Development. In ETS Research Report Series (Vol. 2021, Issue 1, pp. 1–36). Wiley. https://doi.org/10.1002/ets2.12337

Greenberg, N., & Martins, S. J. (2022). Students Perform Faster and Obtain Better Results When Working Together on a Two-Stage Exam. Teaching Tips.
NACTA Journal.

Ian Fellows (2018). wordcloud: Word Clouds. R package version 2.6. https://CRAN.R-project.org/package=wordcloud

Johnson, D. W., & Johnson, R. T. (2009). An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. In Educational Researcher (Vol. 38, Issue 5, pp. 365–379). American Educational Research Association (AERA). https://doi.org/10.3102/0013189x09339057

Le, H., Janssen, J., & Wubbels, T. (2017). Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration. In Cambridge Journal of Education (Vol. 48, Issue 1, pp. 103–122). Informa UK Limited. https://doi.org/10.1080/0305764x.2016.1259389

Linton, D. L., Farmer, J. K., & Peterson, E. (2014). Is Peer Interaction Necessary for Optimal Active Learning? In J. Hewlett (Ed.), CBE—Life Sciences Education (Vol. 13, Issue 2, pp. 243–252). American Society for Cell Biology (ASCB). https://doi.org/10.1187/cbe.13-10-0201

Martins, S. J., Barber, N. L., Bitton-Bailey, A., Byron, M. A., Lyons, N. L., & Roberts, T. G. (2021). Employing Collaborative Assessments in an In-Person Undergraduate Class. NACTA Journal, 65, 478-486.

Molinillo, S., Aguilar-Illescas, R., Anaya-Sánchez, R., & Vallespín-Arán, M. (2018). Exploring the impacts of interactions, social presence and emotional engagement on active collaborative learning in a social web-based environment. In Computers & Education (Vol. 123, pp. 41–52). Elsevier BV. https://doi.org/10.1016/j.compedu.2018.04.012

Penuel, W. R., Krumm, A. E., Pazera, C., Singleton, C., Allen, A., & Deverel-Rico, C. (2023). Belonging in science classrooms: Investigating its relation to students' contributions and influence in knowledge building. In Journal of Research in Science Teaching (Vol. 61, Issue 1, pp. 228–252). Wiley. https://doi.org/10.1002/tea.21884

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

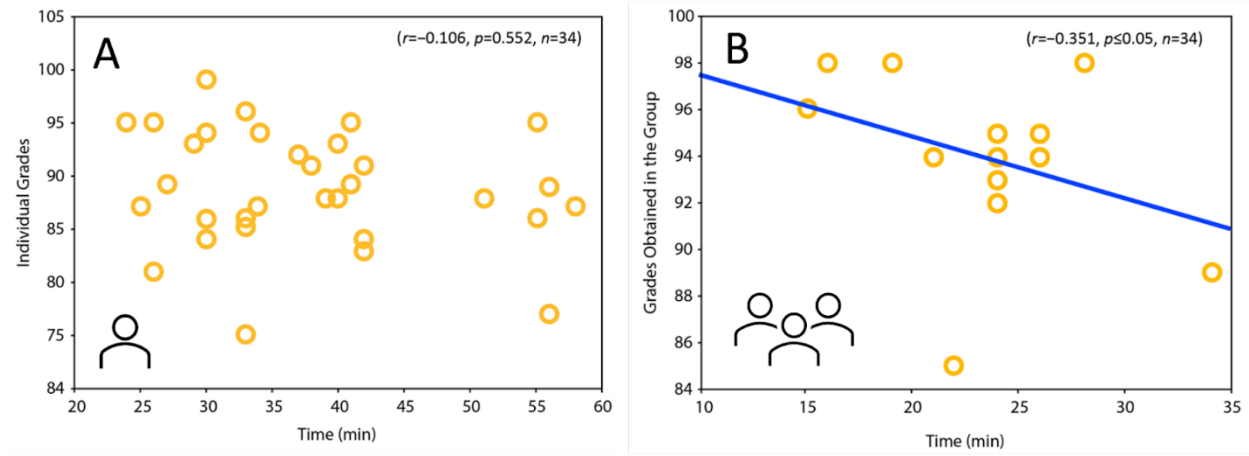
Tullis, J. G., & Goldstone, R. L. (2020). Why does peer instruction benefit student learning? In Cognitive Research: Principles and Implications (Vol. 5, Issue 1). Springer Science and Business Media LLC. https://doi.org/10.1186/s41235-020-00218-5

Turpen, C., & Finkelstein, N. D. (2009). Not all interactive engagement is the same: Variations in physics professors' implementation ofPeer Instruction. In Physical Review Special Topics - Physics Education Research (Vol. 5, Issue 2). American Physical Society (APS). https://doi.org/10.1103/physrevstper.5.020101

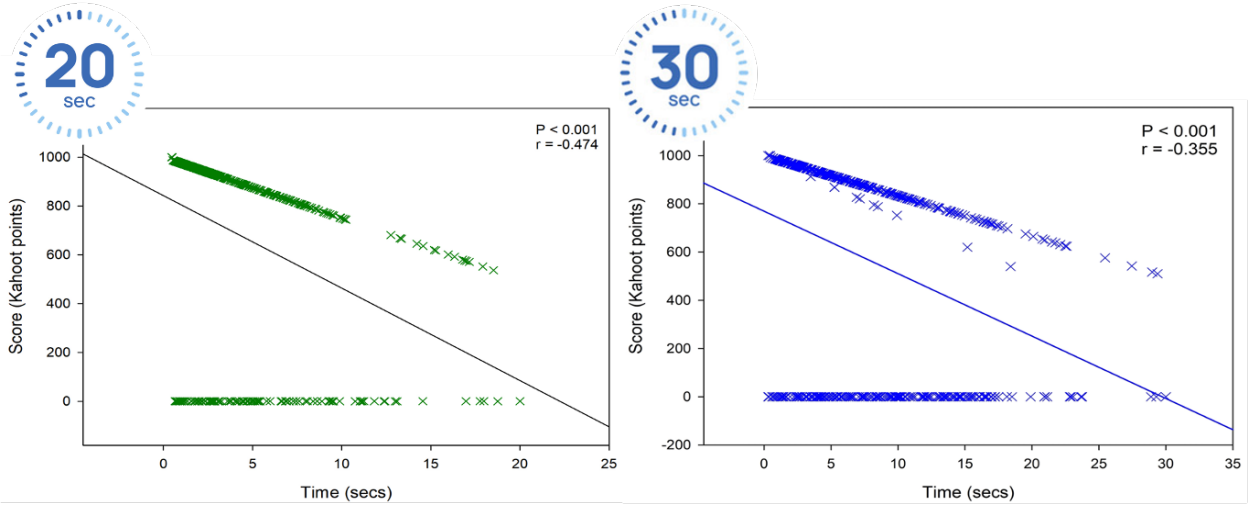
Walton, G. M., & Cohen, G. L. (2011). A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students. In Science (Vol. 331, Issue 6023, pp. 1447–1451). American Association for the Advancement of Science (AAAS). https://doi.org/10.1126/science.1198364

Whitley, E., & Ball, J. (2002). Statistics review 6: Nonparametric methods (Vol. 6, Issue 6, p. 509). Springer Science and Business Media LLC. https://doi.org/10.1186/cc1820

Yeager, D. S., & Walton, G. M. (2011). Social-Psychological Interventions in Education. In Review of Educational Research (Vol. 81, Issue 2, pp. 267–301). American Educational Research Association (AERA). https://doi.org/10.3102/0034654311405999



Supplemental Figure 1 - Pearson's correlation coefficients between students' time to complete the two-stage exam individually (A) and in groups (B) and their individual grades (A) and group grades (B). The data presented represents two exams with seventeen students per exam (n=34).



Supplemental Figure 2 - Time to answer questions vs. score per question for 20 secs (left) and 30 secs (right) when students took the Kahoot! games individually.