Between Integrity and Innovation: STEM Faculty Views on Collaborative Assessment

Allan Jay Dinglasan, Jennifer L. Weible
Central Michigan University
USA

Abstract

Collaborative testing has been shown to enhance student learning and engagement across disciplines, yet adoption in higher education STEM (science, technology, engineering, math) courses remains limited. This qualitative study investigates STEM faculty perceptions of collaborative assessment, with particular attention to issues of academic integrity, individual accountability, and implementation feasibility. Semi-structured interviews were conducted with instructors (n=5) representing biology, chemistry, physiology, statistics, and geosciences who regularly employ collaborative testing in their courses. Thematic analysis identified three key findings. First, while concerns about academic integrity persist, faculty emphasized that integrity issues can be effectively mitigated through structured exam design, clear expectations, and integration of individual components. Second, hesitancy to adopt collaborative testing extends beyond integrity concerns and is deeply rooted in pedagogical beliefs, fears of losing instructional control, and uncertainty about the approach's legitimacy in STEM contexts. Third, successful implementation requires significant effort, including the development of additional materials, grading adjustments, and sustained instructor commitment—efforts that can be prohibitive without institutional support. Despite these barriers, participants argued that collaborative testing promotes accountability, fosters critical thinking, and prepares students for the collaborative demands of the workplace. The findings highlight the need for institutional investment in faculty support, as well as a broader paradigm shift in STEM education toward assessment practices that balance accountability with authentic collaboration. This study contributes to the limited literature on faculty perspectives, underscoring both the challenges and transformative potential of collaborative assessment in higher education STEM disciplines.

1. Introduction

Sir Ken Robinson commented "*When asked about the nature of collaboration"*, how "*in the work world, collaboration and teamwork are essential to success,*" but in school, "it's called cheating." For decades, faculty have often adhered to this, integrating only limited collaborative assessment and group testing across higher education in different disciplines. Although it provided an alternate means of evaluation, with many benefits to students, collaborative assessment also created headaches for faculty with many questions of individual accountability and academic integrity.

The pandemic-induced pivot to remote online learning forced faculty of all disciplines to move to different modalities of assessment and evaluation, often venturing into practices involving group dynamics. In addition, conditions required for education during the pandemic made academic integrity a key point for educators resisting the move to online assessment and evaluation. With the rise of collaborative assessment and group testing, because of its ease of implementation in the online and hybrid learning environment, concerns about academic integrity have tainted its use in the eyes of many faculty, particularly in STEM fields.

Although questions about accountability and integrity are prevalent, limited research exists on collaborative assessment beyond studies of class activities. Therefore, this study seeks to explore experiences of higher education faculty using collaborative assessment in STEM courses and their perceptions of academic integrity, individual accountability, and degree to which the use of collaborative assessment and testing is legitimized and normalized in their field.

2. Literature Review and Conceptual Framework

Studies on the efficacy of collaborative assessment and group work have had mixed results, with the focus being on the challenges that group dynamics pose on collaborative-based forms of assessment, as well as the difficulties differentiating authentic collaboration and cheating. There has been limited research on how academic integrity issues can be mitigated in group settings.

2.1. Challenges of Group Dynamics

Group dynamics has long played a pivotal role in

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

perceptions of the efficacy of any collaboratively inspired pedagogy in colleges and universities. Research has shown that student and instructor perceptions are rooted in preconceived notions and experiences of working as a team in various group settings. Faculty perspectives have implications on the extent of students' entrenched beliefs about how chemistry and science work for the duration of their academic experience [1]. Hamnett et al., found group work as particularly important for postgraduate science students as it provided opportunities for those from diverse cohorts to get familiar with each other's experiences and build their confidence. [2]

Research indicates that strategies for implementation of group work can favor skill development for students through a supportive work environment; this has been shown to promote a more positive attitude as well as perceived greater benefits to their own skill development than their peers who operated in an unsupported work environment [3], although participating in unsupported group activities benefits students by developing their organizational skills.

However, it can be frustrating to work in teams; some members may not fulfill their responsibilities. In particular, higher achieving students fear the possibility of being grouped with those who are less capable [4]. The phenomenon of social loafing is widespread: students often work harder individually than when placed in groups to work collaboratively [5]. Social loafing is prevalent in the workplace, classrooms, and laboratories, as well as learning groups, in both online and face-to-face environments [6].

Likewise, concerns exist about equitable collaboration or groups in which one individual tends to dominate [7]. However, research has found that instructor support of emerging leaders in the course and efforts towards minimizing the problem of free riders can alleviate this issue [8].

2.2. Collaboration vs. Cheating

The division between cheating and collaboration has limited research at this point, focusing more on the mechanics of student cheating. Collaborative assessment has been viewed by all as practices that could create academic integrity issues. In addition, students in higher education STEM courses collaborate much less than those in liberal arts or humanities. Miro Julia and Lopez point to the need to rethink this "collaborative-cheating border", considering faculty concerns of cheating in the computer science discipline.[9] They point to the relativism that is practiced when determining what constitutes cheating: professors' opinions of whether helping a student that is stuck on an individual assignment constitutes cheating or a valid means of collaboration and cooperative learning.

Issues surrounding the assessment of individual performance within a group setting have long triggered resistance from parents and students alike [10]. Additionally, the impact of any potential cheating in that environment is found in anti-cheating measures employed in courses where AI is being used [11]. How faculty and students view what constitutes academic dishonesty and where this collides with genuine collaboration has fueled debate on the merits of collaborative assessments [12], particularly when faculty are asked to differentiate between collaboration and cheating [13]. While prior research had linked cheating with competition, collaboration was also found to have potential for appearance of cheating behaviors with students, something that appears to be culturally specific in nature [14]. Few studies on collaboration versus cheating in academic environments exist [15],[16].

Researchers have proposed adopting collaborative assessment to improve alignment between instruction and assessment, especially in STEM courses preparing students to thrive in the scientific enterprise [17]. In addition, faculty reported having concerns about "unintended collaboration" [18] because of collaborative testing; similar concerns about inflation of grades were reduced by the use of weighted grades [19]. Collaborative assessment, however, has been shown to encompass more than just correct answers. Studies show that more than half of students with incorrect answers prior to collaboration were able to obtain the correct answer following collaboration indicating the positive educational efficacy of collaborative assessments [16].

2.3. Mitigating Academic Integrity Issues

Academic integrity issues have been found to be alleviated with group construction: group size matters with 5-6 being the maximum. Research indicates that smaller groups tend to have less participation [20]. Providing enough time for both individual and group answers allow for more accountability and adequate time to complete both aspects of the test [20]. Newton et al. posits that a grading system with individual and group components, with subsequent weighting allocated by the instructor, promotes individual preparation and rewards for a strong individual performance while simultaneously incentivizing contribution to the group effort [21]. Stearns allowed for two-thirds of the exam grade to come from individual and one third from collaborative efforts [20].

Teacher perspectives on the efficacy of any assessment practice are also fundamental to embracing innovation in this realm. Woodbury and Gess-Newsome proposed a Teacher-Centered Systemic Reform (TCSR) model that highlights the lasting effects of teachers' knowledge and beliefs on their subject matter, pedagogy, and motivation to

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

participate in curriculum reform efforts both within their own classrooms as well as in various institutional contexts [22]. In citing the effect of "paradox of change without difference," the model highlights the importance of alignment with teacher beliefs about content or pedagogy; any reform is bound to fail without it.

2.4. Shifts in Assessment Practices and Faculty "Pedagogical Dissatisfaction"

During the COVID-19 pandemic, many faculty members experienced a sense of "pedagogical dissatisfaction," which created significant pressure to rethink how science courses were delivered in increasingly remote formats [23].

Studies show that while faculty adapted to institutional platforms and resources with relative speed, they often tried to minimize changes to course content. At the same time, they grappled with assessment practices and tensions between their own pedagogical beliefs and the demands of online teaching [24]. Sansom (2020) described this as "deep pedagogical dissatisfaction," as many instructors made instructional choices under pandemic pressures that did not align with their long-held convictions about how chemistry should be taught. Still, history has shown faculty resilience in the face of rapid change: many adapt techniques passed down from previous instructors to strengthen their own practices, fueling a continuous pursuit of teaching improvement [25].

Faculty had to recalibrate curriculum delivery for online contexts, striving to preserve continuity and academic rigor while navigating steep learning curves with limited resources. This led to experimentation with new instructional strategies for lectures, tutorials, group projects, laboratories, and assessments [26].

Responses to this shift were often course specific. While online chemistry homework could be integrated with relative ease, assessment methods demanded greater adjustment and creative approaches [27]. Research points to the value of authentic assessment tasks and activities that foster student self-regulation, while also acknowledging challenges related to academic integrity and faculty hesitancy toward online collaboration—not only in introductory chemistry [28] but across STEM disciplines. As a result, many instructors turned toward alternative assessment formats, including collaborative testing.

2.5. Defining Collaborative Testing

Meseke, Nafziger, and Meseke [29] describe collaborative learning as an active process where students, working in pairs or groups, complete classroom tasks together. Collaborative testing builds on this principle by transforming exams from being solely evaluative to also being opportunities for learning [30].

One common format, the "pyramid" approach, takes place in two stages: first, students complete the exam individually; then, they revisit the same test in groups. Final grades combine individual and group scores [55]. Similarly, Siegel et al. (2015) define group testing as assessments completed collaboratively [17], while two-stage testing requires students to first answer individually, then re-engage with the same assessment in small groups that allow for discussion and reflection beyond traditional test settings [21].

2.6. Documented Student Benefits of Collaborative Testing

Evidence across STEM disciplines demonstrates that collaborative testing supports student learning. Positive outcomes have been reported in biomedical sciences [31], physics [15], physiology [32], biology [33][34], earth and ocean sciences [35], anatomy [36], nursing [37][38], geology [39], astronomy [40], and mathematics [41].

Studies consistently show improved performance when collaborative testing is used [32][39]. Students typically achieve higher scores on group exam components [40], with overall test performance significantly enhanced compared to traditional formats [19]. These improvements occur across question types—including multiple-choice and long-answer assessments [21]—and often foster deeper critical thinking and problem-solving skills [37]. However, such gains may be influenced by factors such as concept complexity, question design, and the duration of content retention [40].

Collaborative testing benefits both high- and low-achieving students [31]. For instance, even strong performers have shown improved accuracy when exams are collaborative [15]. Meanwhile, group scoring often raises outcomes for weaker students [32], [33], [39]. Green [36] cautions, however, that without clear instructional guidance, gains for lower performers may reflect reliance on team scores rather than genuine active learning [36].

2.7. Understanding the Nature of Collaborative Testing Outcomes

The success of collaborative assessments extends beyond the exchange of correct answers; fostering a supportive social environment is essential for sustaining the practice [42], [15]. Research is mixed regarding whether collaborative testing consistently improves retention. Some studies indicate that students retain course content more effectively when testing collaboratively [43], [19], though others found no measurable retention differences despite improved exam performance [33]. Cooke et al. [34] observed

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

stronger retention between pre- and post-tests over longer intervals, but minimal gains over shorter ones. [34] Similarly, Meseke et al. [29] reported improved short-term outcomes but no evidence of enhanced long-term retention.

Collaborative testing also shapes classroom dynamics. Hickey [38] found that students were less confrontational in post-test reviews, as group discussions provided opportunities to raise questions and hear alternative perspectives before seeking clarification from faculty.

2.8. Limited Research on Faculty Perspectives

Empirical studies on faculty perceptions of collaborative testing remain scarce. In one small-scale survey (n=6), instructors felt students learned through the process and valued collaborative testing for accommodating diverse abilities [17]. Concerns over "unintended collaboration" [18] and grade inflation [19] diminished when weighted grading systems were used.

Hickey's [38] small sample (n=4) reported that collaborative testing promoted understanding and critical thinking without reducing preparedness or study habits. Faculty were less certain about effects on test anxiety, ideal group size, or overall study patterns, though they agreed time allotments were adequate.

Research also highlights shifting attitudes among students, with some initially skeptical but later supportive of group exams in mathematics [44]. This raises broader questions about how faculty might foster similar openness toward innovative assessment practices in STEM.

2.9. Conceptual Models and Theoretical Frameworks

Assessment has historically played a minor role in STEM curriculum reform, where emphasis has been placed more on updating content than on rethinking evaluation strategies [45]. For example, while chemistry curricula may incorporate new content such as molecular modeling or statistical applications, assessment often remains rooted in traditional pen-and-paper exams. Bowen and Phelps [45] called for more interactive assessment practices aligned with modern learning environments.

Collaborative assessment fits this vision by integrating learning and evaluation. Rooted in social constructivism, it draws from Johnson and Johnson's [46] social interdependence theory, which frames learning as a product of interaction within social and environmental contexts [47]. Collaborative testing enables students to construct and reconstruct knowledge collectively, while promoting motivation and engagement in innovative classroom settings [48], [49].

Teacher beliefs remain central to adopting new assessment approaches. Woodbury and Gess-Newsome's [22] Teacher-Centered Systemic Reform (TCSR) model emphasizes the interplay between teacher knowledge, pedagogical beliefs, and willingness to embrace reform. Their concept of the "paradox of change without difference" underscores that reform fails when teacher beliefs remain misaligned with proposed changes. Considering these frameworks highlights why many STEM faculty remain hesitant to integrate collaborative assessment into their teaching practices.

3. Methods

This study seeks to examine where instructors are as far as their perceptions of academic integrity and individual accountability within collaborative testing environments in higher education STEM courses. It seeks to add to the limited research on the relationship between faculty perceptions and their hesitance to implement these practices.

3.1. Research Questions

This study seeks to explore the following question: "What are the perceptions of collaborative assessment practitioners in higher education STEM on the role that academic integrity and individual accountability play in an environment where hesitance to implement prevails?"

3.2. Setting and Participants

A virtual conferencing app was used for all interviews to record responses for later analysis and align with best practices for observation of tone and body language [50], which Sturges and Hanrahan found did not change outcomes of the study when using a telephone versus face-to-face interviews. [51] Interviews followed a semi-structured protocol.

A survey was sent to STEM faculty; five (5) participants additionally consented to participate in a semi-structured interview conducted via video conferencing software. These faculty teach STEM courses using a collaborative assessment protocol. Three participants were female and two male from biology, geosciences, statistics, physiology, and chemistry.

3.3. Analysis

Thematic analysis was utilized to draw out emerging patterns using Creswell's process of eliciting codes from interview transcripts, then analyzing and grouping into themes. [52] Merriam and Tisdell describe the constant comparative method where codes are continually compared with one

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

another in a process that triggers more elaborate codes from those found initially. [53] The data were copied into Microsoft Excel where the codes can be more focused and where the themes that emerge can be subsequently grouped.

4. Findings and Discussion

Three main themes emerge from interviews with STEM faculty who are practitioners of collaborative assessment and group testing: 1) Academic integrity issues can be reasonably mitigated, 2) Implementation is rooted in a more deep-seated philosophy regarding pedagogy, and 3) The significant effort required with developing supplementary materials and committing to grade adjustments, that without institutional support can be prohibitive. Additionally, the interviewees point to a need for a paradigm shift to overcome this hesitance to implement an otherwise efficacious practice.

4.1. Mitigating Academic Integrity Concerns

The findings of this study align with research on faculty perceptions of collaborative assessment and group testing that cited concerns about academic integrity but suggested that it is not an issue if a serious effort is made to minimize or alleviate it [18], [19]. Furthermore, they exhibited a keen awareness of its presence in the classroom and have shown considerable willingness to mitigate it. An emphasis on the individual accountability built-in to group assessments as well as having a large percentage of students' grades still being based on individual work is at the core of this mitigative approach.

Instructors faced the academic integrity issues emerging in higher education and specifically from collaborative assessment practices, with realistic expectations that it is an undesirable part of teaching that cannot be easily ameliorated. However, they remain cognizant that putting in an appropriate amount of effort helps moderate instances of perceived cheating.

In Katie's (geosciences) experience, trusting students is something she learned to value in her efforts to maintain academic integrity in her courses. She spoke of the emotional toll and her experience of anger when students do cheat is compounded by what she feels as students taking advantage of their instructors and feeling let down. That feeling turns to fear for Reanne (statistics) who is aware that the preponderance of artificial intelligence (AI) has only made the situation worse. She still prefers to focus on maintaining high engagement levels for her students in the collaborative assessment practices; she feels that academic integrity issues arise when group members become disengaged. Adley (physiology) shared similar awareness of his colleagues' concerns for an increase in *"parasitism"* in the collaborative assessment model, as well as expressing worry that technology used to enhance collaboration may be inadvertently fueling the academic integrity fire: *"To do the collaboration effectively requires.. doing (it) through electronic means…. that also requires the students have access to their laptop throughout... and then there's the whole concern about whether they are just going to Google answers to questions." (Adley)*

Adley uses higher-order thinking questions during the group exams to alleviate students having access to electronic tools and devices during the exam. For Katie, academic integrity issues are minimized when instructors are clear with their expectations for students in so far as what can be shared within the assessments as well as the appropriate way of sharing credit for the responses. Adley added that his group exam protocol has included students being explicitly told that while they can discuss the questions, there cannot be sharing of complete answers. While both the student groups and the instructor took on added responsibility in getting more seemingly unwilling participants to engage in her group exams, Louise (chemistry) believed that having individual components to those exams actually *"skirted the academic integrity concerns."* Gerald (biology) and Adley stated that the conflict between the obligation of institutions of higher learning to elicit individual competence out of their students and the need to prepare them for workplace collaboration certainly needs more attention. Adley took a firmer resolve, saying in part that:

"I certainly realized there can be parasitism (here)…. but I'm not prepared to throw out the whole collaborative model because there's a handful of people abusing the system."

Gerald believes that collaborative assessment allows for individual accountability and has an affordance built-in to the model. Reanne pointed to various *"small things"* that support the existing individual accountability components such as the point distribution system she uses to hold students more accountable to the group. She is convinced that collaborative assessment actually *"fights cheating since a whole group would be very unlikely to cheat."* Louise's students and colleagues have been more receptive to collaborative exams with individual components being written first to demonstrate knowledge. Likewise, Adley stated that incorporating more individual components to a group exam makes collaboration better. Adding to his efforts to build more individual *"nature"* back into his collaborative exams, his students are required to decide for themselves how they are going to answer each individual question. He posited that it did not become a *"group exam"* in the strictest sense of the phrase, since there was only one exam that the whole group

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

submitted. Gerald finds that allocating percentages helps legitimize the practice since some instructors may not think it is fair to give a grade for collaborative assessment. Reanne and Louise are more comfortable with combining the individual part of the exam with the group effort. In her experience, Louise finds combining group and individual scores benefits the students without penalizing them if the group dynamics were not ideal. This adds a *"cushion"* to the individual score that she argued, generates more individual accountability, and helps maintain academic integrity.

4.2. Hesitance Rooted in Pedagogical Beliefs and Uncertainty

Faculty interviewed suggested that colleagues hesitant or dismissive of collaborative assessment practices are also concerned for their own accountability for student learning knowledge and skills required. Despite reassurances that academic integrity issues can indeed be mitigated in this environment, the unanswered questions regarding individual accountability in their roles as course instructors continue to feed hesitance and uncertainty.

Participants pointed also to faculty fearing both a loss of control as a teacher as well as uncertainty regarding the future of collaborative assessment with implementation barriers. This confirmed findings by Marks and O'Connor about the motivations of instructors about any process involving student collaboration [54]. Participants cited lack of confidence and the fear of losing credibility in front of the classroom as reasons driving this hesitance. Katie posited that this fear often drives faculty to embrace pedagogy that is within their *"comfort zone"* when their attempts at trying something new somehow fail:

"The fear is appearing vulnerable, like you don't know what you're doing. Some faculty can recover from making a mistake in classes, and some can't, and…. when there's failure..., they think that they're goanna lose all. credibility." (Katie)

Gerald posited that there are also faculty fears stemming from a mistaken belief that most students are not fans of collaborative assessment. But Reanne agreed that this fear also extends to a loss of autonomy as far as individual assessment goes. *"I think some instructors might be scared to let go of (Individual assessments) ...,and (that)...it's not fair to give a grade for collaborative efforts." (Reanne)*

Uncertainty also appears to drive faculty's hesitance to embrace collaborative assessment practices. Katie spoke of the uncertainty she sees other faculty share regarding whether student understanding was getting truly assessed this way. Adley pointed to not being completely sure that collaborative assessment was helping students, as some have told him that it did, but others would not say otherwise. On the other hand, Louise channeled this same doubt by some veteran educators (on whether collaborative assessment will work) who she said were used to *"doing things a certain way."* She added that they may need to make some adjustments to their professional outlook to make this new pedagogy work for them.

"It does take..courage to try something new, something...you've never experienced as a student…. I don't know how you give people courage beyond trying to get out good research...that this is... important because of the impacts it has on your students and (that) you…should...give it a shot." (Louise)

The instructors also spoke about how the uncertainty surrounding whether collaborative assessment and group testing will ever become part of the traditional evaluation for STEM courses. According to Louise, this lack of optimism is prevalent in higher education circles. She is one of two people actively involved with it in her college. Physiology professor Adley said that rampant skepticism from colleagues regarding collaborative assessment practices is nothing new, despite his numerous individual conversations about it. He said in part that whenever he spoke to his colleagues *"about what I'm doing, most of them would just give (me) that strange look like, "I would never do that."*

4.3. Exerting Effort in Implementation

A third theme that emerges is the perception that while academic integrity concerns about the use of collaborative assessment can indeed be mitigated effectively, it has been their experience that it requires more effort and a willingness on the part of the instructor to prioritize individual accountability within this context. A "buy-in" that occurs on many levels becomes necessary, not just from a philosophical standpoint but from a practical perspective.

In Adley's physiology courses, his experience has been that a lot of extra effort is given by him personally to make sure individual accountability remains paramount, given his colleague's concerns. He found that he had to commit much work to what is normally seen as competing interests in implementing collaborative assessment practices: the need to focus more on actual learning, discussion of ideas, and student community building efforts, as well as that of ensuring that individual accountability concerns are met. He started to incorporate some individual questions into the exams, where in the case of multiple-choice questions, three or four of them within the exam, will be higher order thinking

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

questions that students will get from a pool of questions that he would have had to create or repurpose each time he attempts group testing.
"It makes the collaboration obviously a little harder if everybody's got different questions at that same point in the exam… So, we've tried to build a little bit of individual nature back into some of these collaborative exams." (Adley)

Adley's efforts at ensuring individuals remain accountable in group evaluations has taken time out of getting a broader sense of whether his work is indeed making a difference for his students. While he adheres to the philosophy that what he is doing is ultimately helpful, it would be equally helpful to have the time to constantly revisit and revise:

"I haven't had the opportunity to go and look and see how do students do on those individual questions as opposed to collaborative questions… I haven't had a chance to look and see if it was making any difference there, but I think that would be a way to try and address the concerns...that, students are doing just as well on individual questions as collaborative questions, and perhaps it's not a big deal." (Adley)

In some cases, the instructor's difficulties in implementation, whether it be through the use of certain types of materials or changing up the procedures such as choosing the group composition themselves, (which is time consuming during exams) is superseded by the satisfaction felt from seeing how students react to the way they are assessed. Chemistry instructor Louise adds that fitting more time into her curriculum is worth the effort to make group testing work and students attain the benefits of collaboration:

"I have to give up a fourth hour of class time, instead of just three in order to fit it all in…I think that's a small enough sacrifice to make it work. I don't have any trouble getting through all the content. I have to get through without it…It (does) take a little more time to grade, which is annoying because I have to grade the forms, and then do the calculations, but... grading is a part of teaching. so, I don't think that's a really big detriment to it." (Louise)

Gerald, who had been making the effort to model collaborative assessment and group testing to his biology department, acknowledged that it indeed takes time and effort to implement group testing effectively so that academic integrity remains intact. He cited that a big part of the extra push is in taking time to train students to think more collaboratively. Having a lot of beliefs that are counterintuitive as far as traditional thinking of faculty regarding assessment and evaluation, Steve posits that group testing does gets easier with time, and that when done well, collaborative assessment does not elicit negative reactions from students:

"It is more team, more work and just like, teaching any class or even with collaborative learning kinds of classes…. It gets easier and its less time that's involved, but (when you are still forming groups) at the beginning of the semester. that takes extra time, … and grading the collaborative experiences takes time, but… you can design that in a way that it doesn't take as much time." (Gerald)

On a serious note, there is a recognition in Gerald's experiences that colleagues do have limited time and cannot carve out time to implement collaborative assessment when one considers the need to develop tools to use before attempting to adopt collaborative assessment practices in their classrooms:

"It's more like there's a lot of tools that need to be developed before you can adopt before you can start using it in your class, and so I think it's not that they don't want to, it's just that people's times are limited and, how are you going to carve out the time you have each week to redo your class, so that it can fit in this manner? I think it's an energy issue, how much energy do you have?" (Gerald)

Gerald also cited that the extra effort required is not limited to one's willingness to carve out time to implement collaborative assessment, but also the financial resources to assist the instructors in ensuring that they are not left alone to ensure that no cheating or irregularities occur during group testing.

"That is a big barrier because of the amount of extra grading that goes with this, some faculty are less likely to do this and when (the college) pulled that funding source away and they couldn't have undergrad helpers… you can't get to all of those groups when you need to get to them. So, if you had some experienced undergraduate TAs in the classroom that could filter through the classroom and help those groups, that helps out a lot too, and so I think having extra support in the classroom and with grading would be the thing that most faculty would appreciate having more of." (Gerald)

5. Implications for Practice and Future Research

The findings underscore the need for institutional support structures—such as teaching assistant resources, faculty development programs, and recognition of innovative pedagogy—to reduce the burden of implementation and foster broader adoption of collaborative testing. For practitioners, incorporating individual accountability within group assessments can balance integrity concerns with the

 | DOI: 10.20533/ijcdse.2042.6364.2025.0630

benefits of collaboration. Future research should examine longitudinal outcomes of collaborative testing on student retention, equity, and professional readiness, while also exploring how institutional culture influences faculty willingness to adopt such practices. By situating collaborative testing within a broader framework of systemic support, future scholarship can help STEM education move toward assessment models that are both rigorous and authentically collaborative.

6. Conclusion

The final theme, which highlighted the substantial effort required to design and sustain collaborative testing, provides a crucial link to the conclusion: without adequate time, resources, and institutional support, the potential of collaborative assessment cannot be fully realized. Recognizing the effort-intensive nature of this practice underscores the urgency for systemic change to ensure its viability.

Addressing the legitimate reasons for faculty ambivalence towards collaborative assessment serves the collective interests of all stakeholders in higher education and discourages hesitance towards group testing. It also allows for an attempt at implementation without setting it up for failure in the process. This study found that accountability and academic integrity were foundational in this hesitance. In addition, collaborative assessment has been perceived in the past as faculty ceding control of the evaluation process to the student learners. Faculty utilizing collaborative assessment feel that it makes every stakeholder in the learning process more accountable to each other, while elevating evaluation from beyond a mere check of rote learning to enabling instructors to assess for higher-order thinking skills in the STEM disciplines where students stand to benefit the most.

Practitioners in higher education STEM courses emphasized that while instructors are free to devise their own ways of ensuring accountability and minimizing loafing in collaborative testing, they must also allow for opportunities to work out its inevitable kinks. When faculty outright dismiss collaborative assessment for its flaws without allowing opportunity for improvements, the potential affordances that come with it are greatly diminished, to the detriment of learners and instructors alike.

7. References

[1] Smith, K. C., and Sepulveda, A. (2018). Students' perceptions of common practices, including some academically dishonest practices, in the undergraduate general chemistry classroom laboratory. Chemistry Education Research and Practice, 19(4), 1142–1150. DOI: 10.1039/C8RP00058A.

[2] Hamnett, H. J., McKie, A. E., and Morrison, C. (2018). Postgraduate students' attitudes towards group work: Experiences within a forensic chemistry programme. Chemistry Education Research and Practice, 19(4), 1240–1252. DOI: 10.1039/C8RP00126J.

[3] Natoli, R., Jackling, B., and Seelanatha, L. (2014). The impact of instructor's group management strategies on students' attitudes to group work and generic skill development. Pedagogies, 9(2), 116–132. DOI: 10.1080/1554480X.2014.912519.

[4] Grzimek, V., Marks, M. B., and Kinnamon, E. (2014). Do Differences in GPA Impact Attitudes About Group Work? A Comparison of Business and Non-Business Majors. Journal of Education for Business, 89(5), 263–273. DOI: 10.1080/08832323.2013.872591.

[5] Karau, S. J., and Williams, K. D. (1993). Social Loafing: A meta-analytic review and theoretical integration. Journal of Personality and Social Psychology, 65(4), 681–706. DOI: 10.1037//0022-3514.65.4.681.

[6] Piezon, S. L. (2011). Social loafing and free riding in online learning groups [Ph.D., The Florida State University]. In ProQuest Dissertations and Theses. https://www.proquest.com/docview/897114500/abstract/D120FE02B6D740C7PQ/1 (Access Date: 29 July 2025).

[7] Theobald, E. J., Eddy, S. L., Grunspan, D. Z., Wiggins, B. L., and Crowe, A. J. (2017). Student perception of group dynamics predicts individual performance: Comfort and equity matter. PLOS ONE, 12(7), e0181336. DOI: 10.1371/JOURNAL.PONE.0181336.

[8] Rudawska, A. (2017). Students' Team Project Experiences and Their Attitudes Towards Teamwork. Journal of Management and Business Administration. Central Europe, 25(1), 78–97. DOI: 10.7206/JMBA.CE.2450-7814.190.

[9] Miró Julià, J., and López, D. (2013). Rethinking the collaboration-cheating border. Proceedings of the 18th ACM Conference on Innovation and Technology in Computer Science Education. DOI: 10.1145/2462476.2465603.

[10] King, P. E., and Behnke, R. R. (2005). Problems Associated with Evaluating Student Performance in Groups. College Teaching, 53(2), 57–61.

[11] Goldberg, D. (2021). Programming in a pandemic: Attaining academic integrity in online coding courses. Communications of the Association for Information Systems, 48, 47–54. DOI: 10.17705/1CAIS.04807.

[12] Amzalag, M., Shapira, N., and Dolev, N. (2021). Two sides of the coin: Lack of academic integrity in exams during the Corona pandemic, students' and lecturers' perceptions. Journal of Academic Ethics, 20(2), 243–263. DOI: 10.1007/S10805-021-09413-5.

[13] Deale, C. S., (Jenna) Lee, S. H., (Stephanie) Bae, J.-I., and White, B. J. (2019). An exploratory study of educators' and students' perceptions of collaboration versus cheating in Hospitality and Tourism Education. Journal of Teaching

in Travel and Tourism, 20(2), 89–104. DOI: 10.1080/15313220.2019.1679695.

[14] Jamaluddin, S. F., Adi, S. P., and Lufityanto, G. (2021). Social influences on cheating in Collectivistic Culture: Collaboration but not competition. Group Dynamics: Theory, Research, and Practice, 25(2), 174–189. DOI: 10.1037/GDN0000122.

[15] Jang, H., Lasry, N., Miller, K., and Mazur, E. (2017). Collaborative exams: Cheating? Or learning? American Journal of Physics, 85(3), 223–227. DOI: 10.1119/1.4974744.

[16] Sheard, J., Dick, M., Markham, S., Macdonald, I., and Walsh, M. (2002a). Cheating and plagiarism: perceptions and practices of first year IT students. Proceedings of the 7th Annual Conference on Innovation and Technology in Computer Science Education. DOI: 10.1145/544414.544468.

[17] Siegel, M., Roberts, T., Freyermuth, S., Witzig, S., and Izci, K. (2015). Research and teaching: Aligning assessment to instruction: Collaborative group testing in large-enrollment science classes. Journal of College Science Teaching, 044(06). DOI: 10.2505/4/JCST15_044_06_74.

[18] Lewis, S. E. (2020). Chemistry assessments through the sudden implementation of online instruction. Journal of Chemical Education, 97(9), 3418–3422. DOI: 10.1021/ACS.JCHEMED.0C00697.

[19] Bloom, D. (2009). Collaborative test taking: Benefits for learning and retention. College Teaching, 57(4), 216–220. DOI: 10.1080/87567550903218646.

[20] Stearns, S. A. (1996). Collaborative Exams as Learning Tools. College Teaching, 44(3), 111–112.

[21] Newton, G., Ritchie, K., Albabish, W., Gilley, B., Rajakaruna, R., and Kulak, V. (2019). Research and teaching: Two-stage (collaborative) testing in science teaching: Does it improve grades on short-answer questions and retention of material? Journal of College Science Teaching, 048(04). DOI: 10.2505/4/JCST19_048_04_64.

[22] Woodbury, S., and Gess-Newsome, J. (2002). Overcoming the Paradox of Change without Difference: A Model of Change in the Arena of Fundamental School Reform. Educational Policy, 16(5), 763–782. DOI: 10.1177/089590402237312.

[23] Sansom, R. L. (2020). Pressure from the pandemic: Pedagogical dissatisfaction reveals faculty beliefs. Journal of Chemical Education, 97(9), 2378-2382. DOI: 10.1021/ACS.JCHEMED.0C00657.

[24] Rupnow, R. L., LaDue, N. D., James, N. M., and Bergan-Roller, H. E. (2020). A perturbed system: How tenured faculty responded to the COVID-19 shift to remote instruction. Journal of Chemical Education, 97(9), 2397–2407. DOI: 10.1021/ACS.JCHEMED.0C00802.

[25] Kusch, J. (2016). Learning experiences of university biology faculty: A qualitative pilot study. Journal of College Science Teaching, 045(06). DOI:10.2505/4/JCST16_045_06_24.

[26] Dietrich, N., Kentheswaran, K., Ahmadi, A., Teychené, J., Bessière, Y., Alfenore, S., Laborie, S., Bastoul, D., Loubière, K., Guigui, C., Sperandio, M., Barna, L., Paul, E., Cabassud, C., Liné, A., and Hébrard, G. (2020). Attempts, successes, and failures of distance learning in the time of covid-19. Journal of Chemical Education, 97(9), 2448–2457. DOI: 10.1021/ACS.JCHEMED.0C00717.

[27] Clark, T. M., Callam, C. S., Paul, N. M., Stoltzfus, M. W., and Turner, D. (2020). Testing in the time of covid-19: A sudden transition to unproctored online exams. Journal of Chemical Education, 97(9), 3413–3417. DOI: 10.1021/ACS.JCHEMED.0C00546.

[28] Raje, S., and Stitzel, S. (2020). Strategies for effective assessments while ensuring academic integrity in general chemistry courses during COVID-19. Journal of Chemical Education, 97(9), 3436–3440. DOI: 10.1021/ACS.JCHEMED.0C00797.

[29] Meseke, C. A., Nafziger, R., and Meseke, J. K. (2010). Student attitudes, satisfaction, and learning in a collaborative testing environment. Journal of Chiropractic Education, 24(1), 19–29. DOI: 10.7899/1042-5055-24.1.19.

[30] Muir, S. P., and Tracy, D. M. (1999). Collaborative essay testing: Just try it! College Teaching, 47(1), 33–35. DOI: 10.1080/87567559909596077.

[31] Khong, M. L., and Tanner, J. A. (2020). A collaborative two-stage examination in BIOMEDICAL sciences: Positive impact on feedback and peer collaboration. Biochemistry and Molecular Biology Education, 49(1), 69–79. DOI: 10.1002/BMB.21392.

[32] Giuliodori, M. J., Lujan, H., and DiCarlo, S. E. (2008). Collaborative group testing benefits high performing as well as low performing students. The FASEB Journal, 22(S1). DOI: 10.1096/FASEBJ.22.1_SUPPLEMENT.766.4.

[33] Leight, H., Saunders, C., Calkins, R., and Withers, M. (2012). Collaborative testing improves performance but not content retention in a large-enrolment introductory biology class. CBE—Life Sciences Education, 11(4), 392–401. DOI: 10.1187/CBE.12-04-0048.

[34] Cooke, J. E., Weir, L., and Clarkston, B. (2019). Retention following two-stage collaborative exams depends on timing and student performance. CBE—Life Sciences Education, 18(2). DOI: 10.1187/CBE.17-07-0137.

[35] Gilley, B., and Clarkston, B. (2014). Research and teaching: Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. Journal of College Science Teaching, 043(03). DOI: 10.2505/4/JCST14_043_03_83.

[36] Green, R. A., Cates, T., White, L., and Farchione, D. (2015). Do collaborative practical tests encourage student-centered active learning of gross anatomy? Anatomical Sciences Education, 9(3), 231–237. DOI: 10.1002/ASE.1564.

[37] Rivaz, M., Momennasab, M., and Shokrollahi, P. (2015, October 1). Effect of collaborative testing on learning and retention of course content in nursing students. Journal of Advances in Medical Education and Professionalism. Retrieved September 20, 2021, from http://jamp.sums.ac.ir/index.php/JAMP/article/view/354 (Access date: 29 July 2025).

[38] Hickey, B. L. (2006). Lessons learned from collaborative testing. Nurse Educator, 31(2), 88–91. DOI: 10.1097/00006223-200603000-00012.

[39] Knierim, K., Turner, H., and Davis, R. K. (2015). Two-stage exams improve student learning in an introductory geology course: Logistics, attendance, and grades. Journal of Geoscience Education, 63(2), 157–164. DOI: 10.5408/14-051.1.

[40] Miller, S. T. J. (2019). What is the impact of collaborative exams on learning and attitudes in introductory astronomy classes? Journal of Astronomy and Earth Sciences Education. https://eric.ed.gov/?id=EJ12 2 8738 (Access date: 20 September 2021).

[41] Kinnear, G. (2020). Two-stage collaborative exams have little impact on subsequent exam performance in undergraduate mathematics. International Journal of Research in Undergraduate Mathematics Education, 7(1), 33–60. DOI: 10.1007/s40753-020-00121-w.

[42] McConnell, D. (2002). The experience of collaborative assessment in e-learning. Studies in Continuing Education, 24(1), 73–92. DOI: 10.1080/01580370220130459.

[43] 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. Advances in Physiology Education, 27(3), 102–108. DOI: 10.1152/ADVAN.00041.2002.

[44] Dobie, T. E., and MacArthur, K. (2022). Exploring Shifts in Student Attitudes Toward Group Exams in College Calculus: The Case of Dane. International Journal of Research in Undergraduate Mathematics Education, 8(1), 149–175. DOI: 10.1007/S40753-021-00148-7.

[45] Bowen, C. W., and Phelps, A. J. (1997). Demonstration-Based Cooperative Testing in General Chemistry: A Broader Assessment-of-Learning Technique. Journal of Chemical Education, 74(6), 715. DOI: 10.1021/ED074P715.

[46] Johnson, D. W., and Johnson, R. T. (2015). Cooperation and Competition. DOI: 10.1016/B978-0-08-097086-8.24051-8.

[47] Mishra, R. K. (2014). Social Constructivism and Teaching of Social Science. Journal of Social Studies Education Research, 5. DOI: 10.17499/JSSER.22283.

[48] American Association for the Advancement of Science (2011). Vision and change in introductory biology education: A call to action. Washington, DC.

[49] National Research Council (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press.

[50] Opdenakker, Raymond (2006, August). Advantages and Disadvantages of Four Interview Techniques in Qualitative Research [44 paragraphs]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 7(4), Art. 11, http://nbn-resolving.de/urn:nbn:de:0114-fqs0604118 (Access Date: 29 July 2025).

[51] Sturges, J. and Hanrahan, K. (2004). Comparing Telephone and Face-to-Face Qualitative Interviewing: a Research Note. Qualitative Research. 4(1):107-118. DOI: 10.1177/1468794104041110.

[52] Creswell, J. W. (2014). Educational research: Planning, conducting and evaluating quantitative and qualitative research. Pearson.

[53] Merriam, S. B., and Tisdell, E. J. (2017). Qualitative research: A guide to design and implementation. Langara College.

[54] Marks, M. B., and O'Connor, A. H. (2013). Understanding Students' Attitudes About Group Work: What Does This Suggest for Instructors of Business? Journal of Education for Business, 88(3), 147–158. DOI: 10.1080/08832323.2012.664579.

[55] Zipp, J. F. (2007). Learning by exams: The impact of two-stage cooperative tests. Teaching Sociology, 35(1), 62–76. DOI: 10.1177/0092055X0703500105.

