International Journal of Research in Education Humanities and Commerce

Volume 07, Issue 05 "September - October 2026"

ISSN 2583-0333

ENHANCING THE EVALUATION OF GROUP-BASED ASSIGNMENTS THROUGH STRUCTURED PEER ASSESSMENT AMONG PRE-SERVICE TEACHERS AT BISHOP STUART UNIVERSITY

MWEBEMBEZI ISRAEL[1], ENOCK BARIGYE, PhD[2], BASIL TIBANYENDERA, PhD[3] & Rev. Dr. JUDITH ARINAITWE[4]

[1], [2], [3], [4] Bishop Stuart University, Mbarara, Uganda

[1]Master of Education Management, Planning and Administration (MEMPA) Candidate; [2],[3],[4]PhD

https://doi.org/10.37602/IJREHC.2026.7514

ABSTRACT

This study explored how structured peer assessment could improve the evaluation of group-based assignments among pre-service teachers at Bishop Stuart University, guided by Equity Theory (Adams, 1965). A qualitative, evaluative action research design used a Focus Group Discussion Guide, a Coordinator Interview Guide, a Classroom Observation Schedule, a Document Review Checklist, and a Structured Peer Assessment Tool, generating data from 540 students across 55 groups. Findings showed that group assignments were previously graded uniformly regardless of individual input, tolerating free-riding. The intervention, implemented through individual and private rating, produced genuine differentiation in every group, averaging seven points of spread out of 100, with increased individual accountability. Challenges included reciprocal marking, collusion, and group-size effects. The study recommends smaller, deliberately composed groups and institutional policy development to formalise structured peer assessment.

Keywords: peer assessment, group-based assignments, equity theory, contribution-weighted grading, higher education, Uganda

1.0 INTRODUCTION

Group-based assignments are a standard feature of higher education, valued for developing the collaborative and interpersonal competencies that individual assessment cannot adequately test, and for allowing institutions to manage growing student populations without proportionally increasing assessment workloads (Boud, Ajjawi, Dawson, & Tai, 2022). Yet a persistent problem undermines their fairness: when every member of a group receives an identical mark regardless of individual input, the assessment fails to reward the students who actually did the work, and structurally tolerates free-riding by those who did not (Kaufman, Felder, & Fuller, 2000; Simms & Nichols, 2014). At Bishop Stuart University, this uniform grading practice has remained the dominant norm in the Faculty of Education, Arts and Media Studies, despite the availability of structured peer assessment as an alternative that has been shown elsewhere to differentiate grades according to actual contribution (Topping, 2018; Heng, Chu, & Kong, 2020).

 Copyright © The Author, All rights reserved

This study was guided by Equity Theory (Adams, 1965), which holds that individuals compare the ratio of their inputs, such as effort and contribution, to their outcomes, such as rewards and grades, against that of others, and experience resentment or reduced effort when that ratio is perceived as unfair. Uniform grading, by holding outcomes constant while inputs vary widely across group members, is a direct structural cause of the inequity Adams describes: it guarantees that most students in any realistically varied group receive a grade that does not match what they actually contributed. The purpose of this study was to explore how structured peer assessment could be used to improve the evaluation of group-based assignments among pre-service teachers at Bishop Stuart University. Four subsidiary questions guided the study: (1) how group-based assignments were evaluated before the introduction of structured peer assessment; (2) what procedures could be used to implement structured peer assessment; (3) how the evaluation of group-based assignments changed after its introduction; and (4) what challenges arose in using structured peer assessment. Addressing these questions is significant both practically, in offering Bishop Stuart University an evidence-based route to fairer group assessment and a contribution toward its institutional assessment policy, and theoretically, in testing Equity Theory's predictions within an under-studied East African higher education setting.

2.0 LITERATURE REVIEW

2.1 Group Assignment Assessment and the Uniform Grading Problem

Group-based assignment became standard practice in higher education for reasons that seemed both pedagogically sound and practically convenient, but the relationship between group performance and individual grades was less carefully considered in the early decades of its expansion. The default arrangement, a single collective score applied equally to every member, treats the group as an indivisible unit and ignores the reality that group members contribute at very different rates. Falchikov (2005) traced the literature on “social loafing”, the tendency for group members to reduce effort when a grade will be equally shared, and identified the uniform grading model as its primary structural cause. Oakley, Felder, Brent, and Elhajj (2004) argued further that it is not merely that some students choose to disengage, but that the assessment system actively invites disengagement by removing any grade-related incentive for individual effort. Double, McGrane, and Hopfenbeck (2020), synthesising decades of subsequent research, confirmed that this pattern holds across disciplines, institutional types, and national contexts: equal-grade group assessment produces predictably unequal contribution, with high-contributing students bearing a cost they did not choose and low-contributing students enjoying a benefit they did not earn.

This pattern has been particularly under-addressed in Sub-Saharan Africa. Ugandan universities adopted group-based assessment largely as a pragmatic response to under-resourcing and overcrowding, without simultaneously developing tools to ensure individual contributions were recognised (Boud & Falchikov, 2007). Muzaale, Kasujja, and Kasekende (2026), studying assessment practices across Ugandan universities, found uniform grading so embedded as the unquestioned norm that neither policy frameworks nor practitioner protocols made any provision for differentiating individual contribution within group marks.

2.2 Peer Assessment as a Mechanism for Fairer Evaluation

 Copyright © The Author, All rights reserved

Volume 07, Issue 05 "September - October 2026"

ISSN 2583-0333

Peer assessment, the process through which students evaluate and rate one another's contributions against shared, predefined criteria, has been studied since the late 1970s. Early, informal implementations produced mixed results: without clear rating criteria, ratings were inconsistent; without anonymity protections, friendship and social dynamics distorted judgement; and where peer scores carried no weight in the final grade, students had little reason to take the exercise seriously. These early failures clarified what a well-designed peer assessment system needed. Cheng and Warren (2000) demonstrated that criterion clarity, the degree to which rating dimensions are explicit and tied to observable behaviour, is the strongest single predictor of peer rating accuracy. Wen and Tsai (2006) showed that students who received structured orientation before rating produced ratings significantly more consistent and reliable than those who did not. Falchikov (2005) identified anonymity as the most effective structural protection against friendship bias, leniency, and strategic inflation, a finding Double et al. (2020) confirmed in a meta-analysis showing that anonymous peer assessment with explicit behavioural criteria produced the strongest gains in rating accuracy. Beyond producing accurate contribution scores, Topping (2018) found that peer assessment develops evaluative judgement and critical thinking in the raters themselves, skills directly relevant to pre-service teachers who will later design and administer their own assessments. Liu and Carless (2006) found that structured peer assessment reduced unequal contribution over time as students adjusted behaviour in response to knowing peer ratings would count, a motivational pattern consistent with Equity Theory's prediction that free-riding declines once the contribution-to-grade link becomes transparent. Kaufman, Felder, and Fuller (2000) confirmed this empirically, finding that contribution-weighted grading produced more proportionate effort distribution than uniform grading.

2.3 Peer Assessment and Grade-to-Contribution Alignment

Under uniform grading, every student who contributes at a rate other than exactly the group average, which in any realistically varied group means most students, receives a grade that does not match their actual contribution. Simms and Nichols (2014), reviewing seventy-eight studies on group work in higher education, found that equal-grading contexts consistently produced the lowest grade-to-contribution accuracy, with high-contributing students the most severely disadvantaged. Structured peer assessment changes this picture. Heng, Chu, and Kong (2020) found that students who participated in structured peer contribution assessment received final grades that measurably differentiated between high and low contributors, closely tracking independent assessments of actual contribution. Kaufman et al. (2000) demonstrated that contribution-weighted grading produced a statistically significant spread of individual grades proportional to peer-rated contribution, in contrast to the flat distribution produced under equal-mark grading. In the most comprehensive synthesis to date, Panadero, Jonsson, and Botella (2019) analysed forty peer assessment studies across disciplines and confirmed that structured peer assessment consistently improved grade-to-contribution alignment across science, engineering, humanities, and social science cohorts, and across universities in Europe, North America, and Asia. The Ugandan evidence base is considerably thinner: Muzaale et al. (2026) documented that grade differentiation relative to contribution is effectively absent across most Ugandan university group assessment contexts, establishing the baseline this study sought to change. No published study had yet evaluated a structured peer assessment intervention at Bishop Stuart University, confirming both the originality and the practical urgency of the present study.

 Copyright © The Author, All rights reserved

Dochy, Segers, and Sluijsmans (1999) first established the link between well-designed peer assessment and grade-to-contribution accuracy in the 1990s; Topping (2018) refined it by identifying the specific design features most reliably associated with accurate alignment; and Panadero et al. (2019) and Double et al. (2020) confirmed it in meta-analyses spanning more studies and contexts than any single investigation could reach. What the international literature does not provide is a locally situated account of how such an intervention unfolds within a Ugandan university, how students and staff experience it, and what specifically changes in the grade-to-contribution relationship within this particular institutional and cultural context. Herr and Anderson (2022) note that practitioner-led interventions are especially effective at producing locally relevant improvement, because the researcher's contextual knowledge allows the intervention to be calibrated to barriers an outsider would not identify. This study was designed to generate exactly this kind of locally grounded, immediately actionable evidence, addressing a gap that is both empirical and practical for assessment policy at Bishop Stuart University.

3.0 METHODOLOGY

This study adopted a qualitative, evaluative action research design (Kemmis, McTaggart, & Nixon, 2004), selected because the research problem required both a practical intervention and a rigorous evaluation of that intervention's effect within a real institutional setting, rather than recommendations produced from outside it. The design unfolded across three phases corresponding to the plan-act-observe-reflect cycle: a pre-intervention phase establishing the baseline evaluation practice and diagnostic information for calibrating the intervention; an intervention phase implementing the structured peer assessment process; and a post-intervention phase evaluating its effect and the challenges it generated. Equity Theory (Adams, 1965) served as the interpretive lens throughout, providing the standard, grades proportional to contribution, against which each phase's data was read.

The study population comprised pre-service teachers in the foundational course units of the Faculty of Education, Arts and Media Studies, from which a purposive, maximum-variation sample of sixteen focus students was drawn (four per department, anonymised in reporting as Departments A, B, C, and D), together with two programme coordinators selected purposively for their oversight of assessment practice. Data were collected using five principal instruments. A Focus Group Discussion Guide was administered across eight FGDs, four pre-intervention and four post-interventions, each stratified by department, to capture participants lived experience of assessment before and after the intervention. A structured Coordinator Interview Guide was administered before and after the intervention to capture institutional-level perspective and barriers. A Classroom Observation Schedule was used throughout the orientation and implementation of the peer assessment tool to record procedural detail that interview and FGD data alone could not capture. A Document Review Checklist was applied to departmental coursework and grading records to establish the formal, documented baseline practice. Finally, a Structured Peer Assessment Tool generated quantitative grade data for 540 students across 55 groups.

The structured peer assessment intervention organised students into groups of ten, a size chosen to give each member enough peer raters to generate stable, representative contribution scores while remaining small enough for individual contribution to stay visible to the group (Kaufman

 Copyright © The Author, All rights reserved

Volume 07, Issue 05 "September - October 2026"

ISSN 2583-0333

et al., 2000). Implementation proceeded through four sequential steps. First, a structured orientation session introduced the purpose of peer assessment, the rating dimensions, the anonymity protocol, and a worked example, reflecting Wen and Tsai's (2006) finding that orientation quality is among the strongest predictors of rating accuracy. Second, students were organised into groups of exactly ten for a shared assignment. Third, on completion of the assignment, each student rated the other nine members individually and privately, away from the group, on four dimensions, Commitment, Quality of Work Contributed, Collaboration, and Adherence to Timelines, using a five-point scale; this anonymity and privacy reflects Falchikov's (2005) identification of anonymity as the strongest structural protection against friendship bias and retaliation. Fourth, a contribution-weighting factor, derived from each student's average peer score relative to their group's mean peer score, was applied to the group's collectively awarded assignment score to compute each student's final, individually differentiated grade.

Qualitative data were analysed thematically following Braun and Clarke's (2022) six-phase approach: familiarisation with the full data corpus; generation of initial codes; construction of candidate themes; review of themes against the raw data; definition and naming of final themes; and interpretation against Equity Theory as the standing analytic lens. Quantitative grade data were analysed descriptively, computing within-group score ranges before and after the intervention across all 55 groups. Ethical clearance was obtained from the Bishop Stuart University Research Ethics Committee prior to data collection. Participation in all activities was voluntary, informed consent was obtained in writing after the Informed Consent Form was read through with each participant, and confidentiality was protected through pseudonymisation, anonymised departmental labelling, and sealed, unidentifiable submission of peer rating rubrics.

Trustworthiness was pursued through triangulation across the five data sources: qualitative claims were checked against quantitative grade patterns, and vice versa, and consistency was sought across independent accounts from students and coordinators before a theme was treated as established. A reflective journal and periodic debriefing with a collaborating colleague were maintained throughout the three phases to surface researcher assumptions and challenge emerging interpretations, consistent with the reflective requirements of action research (Kemmis et al., 2004). Several limitations should nonetheless be noted. The quantitative grade data reflects a single implementation cycle, so it is not yet possible to establish whether the differentiation observed would persist across repeated use. The purposive focus-student sample, while stratified across four departments, remains modest in size relative to the wider student population, and the findings are most directly generalisable to comparable large-enrolment, foundational-course contexts rather than to Bishop Stuart University as a whole.

4.0 FINDINGS

4.1 Pre-Intervention Evaluation Practice

Across all four pre-intervention FGDs, participants consistently reported that group assignments were graded through a single, uniform mark applied to every member regardless of individual input, with students describing identical scores such as “80-80” or “20-19-19” irrespective of who had done the work. This was corroborated independently by the programme coordinators and by the document review of departmental coursework records, which showed

 Copyright © The Author, All rights reserved

Volume 07, Issue 05 "September - October 2026"

ISSN 2583-0333

the same uniform grading pattern across departments, with only a few course units adding presentation marks.

> *"So now the problem is where we are given the same mark in the group even when we have not contributed." (FGD Participant, Department B, Pre-Intervention)*

Where differentiation did occur, participants pointed to individual presentation as the main, and often only, mechanism, but viewed it as unreliable in classes exceeding 500 students, where lecturers could not feasibly observe every member present, and where some students simply read prepared content aloud without demonstrating real understanding.

Free-riding was tolerated under uniform grading: some members contributed money for printing or logistics rather than intellectual effort, and were still rewarded identically. Participants explained that non-contributing members effectively treated the few students who did the real work as their "academic houseboys", diligent students working, unpaid, on behalf of others who then claimed an equal share of the reward without having contributed to earning it.

> *"He can just give out the money for printing and he just comes to sign... everybody contributes like 1,000 shillings each and gets his 10,000 and makes it a business, maybe." (FGD Participant, Department A, Pre-Intervention)*

Group leaders reported deliberately avoiding conflict by not excluding non-contributing members, for fear of confrontation, allowing free-riding to persist unchecked. The coordinators independently confirmed no formal peer assessment policy existed at Bishop Stuart University prior to the intervention.

4.2 Implementation Procedures

The core design decision, individual and private completion of peer ratings, was affirmed by coordinators and students alike as central to honest rating, since it removed the immediate social cost of giving an unfavourable score. Students reflecting after the fact confirmed this directly.

> *"When I was individual, on my bed or somewhere, it gave me that confidence that I'm giving this person this credit, because this is what he or she performed." (FGD Participant, Department B, Post-Intervention)*

Considerable orientation effort was required, addressing student questions about identifiers and anonymity, and technical accommodations, including shared device use and extended submission windows, supported participation among students without functioning phones. Some groups nonetheless attempted to coordinate uniform ratings in discussion before submission, effectively trying to reproduce the old uniform-grading pattern within the new tool; facilitator intervention, reminding students that contribution levels genuinely differ, was needed to shift some submissions toward more genuine rating.

4.3 Post-Intervention Change

 Copyright © The Author, All rights reserved

Volume 07, Issue 05 "September - October 2026"

ISSN 2583-0333

Quantitative analysis of the structured peer assessment tool data, 540 students across 55 groups, showed that differentiation, entirely absent under uniform grading, emerged consistently once peer-informed weighting was applied. Table 1 summarises the key figures.

Indicator	Value
Total students / groups	540 / 55
Mean pre-intervention group score (/100)	74.51
Mean post-intervention individual score (/100)	61.70
Mean within-group differentiation range	6.97 points
Groups with range > 10 points	9 (16.4%)
Groups with range ≤ 3 points	5 (9.1%)

Table 1: Summary of Pre- and Post-Intervention Grade Differentiation

Every one of the 55 groups showed some degree of differentiation once peer ratings were applied, with an average spread of roughly seven points out of 100 between the highest- and lowest-scoring member of a group. Nine groups (16.4%) showed substantial differentiation exceeding ten points, while a smaller cluster of five groups (9.1%) showed a narrow spread of three points or less, which may reflect either genuinely comparable contribution or unaddressed leniency bias. This pattern was corroborated qualitatively: participants described a perceptible grading gradient from higher to lower contributors.

> *"No, it did not keep the same. That difference emerged... there is a gradual increase or decrease from the top person to the one which is the lowest." (FGD Participant, Department B, Post-Intervention)*

Both students and coordinators also reported increased individual accountability: because students knew their peers would rate their actual contribution, several described preparing materials they had previously neglected, in anticipation of peer evaluation.

> *"If you know that you have nothing to contribute on the work... today evening you go and read, possibly also get some credit tomorrow." (FGD Participant, Department A, Post-Intervention)*

4.4 Challenges

The intervention surfaced reciprocal and retaliatory marking, where students pre-emptively rated a peer poorly in anticipation of receiving a poor rating themselves, regardless of that peer's actual contribution.

> *"I know Mr Israel will give me a poor... let me also give him or her a poor, because I know he will not give me much." (FGD Participant, Department D, Post-Intervention)*

A more specific and, in the researchers' assessment, more serious pattern was also identified: a majority-coalition bias, in which non-contributing members who numerically outnumbered genuine contributors within a group could collectively out-rate that minority, disadvantaging the very students the intervention was designed to protect. Attempted collusion around uniform

 Copyright © The Author, All rights reserved

scores, agreed in advance by some groups, was also documented, alongside difficulties associated with large, friendship-based groups that made individual monitoring difficult and encouraged protective rating. Coordinators recommended reducing group sizes from fifteen to twenty members to approximately ten, and forming groups based on seriousness and willingness to contribute rather than friendship. Finally, uneven initial buy-in was documented: some students treated the exercise as a low-stakes activity, almost a game, until final, differentiated marks were released and its real consequences became apparent.

> *"Do you know that I contributed much in this work?... how come I ended up having 14 out of 20?" (FGD Participant, Department C, Post-Intervention)*

5.0 DISCUSSION

The uniform grading practice documented here is a direct illustration of the inequity Adams (1965) described: outcomes held constant while inputs varied widely, producing a perceived imbalance in the input-outcome ratio that participants articulated unprompted through language such as "academic houseboys." This agrees closely with Kaufman, Felder, and Fuller (2000) and Simms and Nichols (2014), who argue that uniform grading not only fails to prevent free-riding but structurally rewards it by making group membership, rather than group contribution, the basis for reward. Where this study extends the existing literature is in documenting conflict-avoidance as the specific mechanism sustaining tolerance for free-riding: group leaders did not report indifference to non-contribution, but active, conscious tolerance of it, driven by fear of confrontation, a dimension not fully captured in the largely North American accounts this literature is otherwise drawn from.

The success of individual, private rating in producing honest differentiation agrees with Falchikov's (2005) argument that anonymity is a precondition, not merely an enhancement, for honest peer rating, since raters who know their scores will be visible to, or negotiated with, the person being rated tend to converge toward leniency regardless of actual contribution. The extensive orientation effort documented in this study reflects Wen and Tsai's (2006) finding that rating accuracy depends heavily on students understanding the purpose and stakes of the exercise beforehand, while the technical accommodations made, extended submission windows and permitted device-sharing, reflect Double, McGrane, and Hopfenbeck's (2020) observation that practical usability, not only instrument design, determines whether students engage seriously. The early resistance observed, groups attempting to coordinate uniform ratings before submission, is a useful corrective to any assumption that a well-designed private instrument automatically produces independent judgement: anonymity of submission and independence of judgement appear to be two distinct properties that must be secured separately.

The consistent emergence of differentiation across all 55 groups agrees with Heng, Chu, and Kong (2020) and supports Kaufman et al.'s (2000) broader claim that contribution-weighting mechanisms work specifically by making previously invisible effort differences visible in the grade itself. The accompanying behavioural evidence, students reporting increased preparation once they knew peers would rate them, extends Equity Theory's predictions from a purely evaluative claim toward a motivational one, consistent with Topping's (2018) argument that peer assessment's value lies as much in its effect on the collaborative process itself as in the accuracy of its final scores, and with Dochy, Segers, and Sluijsmans's (1999) observation that

 Copyright © The Author, All rights reserved

positioning students as active evaluators of one another tends to increase engagement independent of the resulting grades.

The challenges identified, particularly reciprocal marking and collusion, agree with Falchikov (2005), Liu and Carless (2006), and Cheng and Warren (2000), who caution that peer assessment can reproduce, rather than correct, existing social dynamics within a group whenever raters are not genuinely independent of one another. The majority-coalition bias identified in this study is a more specific and practically serious variant of this risk than is typically emphasised in the literature, which tends to discuss reciprocity bias at the level of individual rater pairs rather than group-level voting dynamics; this is a finding this study contributes to the field rather than one purely confirmed by it, and it suggests that simple averaging of peer ratings may be insufficiently robust wherever non-contributors are numerically dominant within a group. Consistent with Oakley, Felder, Brent, and Elhajj's (2004) guidance on deliberate group formation, sustaining the gains observed here would require ongoing orientation, appropriately sized and deliberately composed groups, and moderation mechanisms such as flagging statistically extreme outlier ratings, rather than treating the intervention as complete after a single implementation cycle.

6.0 CONCLUSIONS AND RECOMMENDATIONS

This study concludes that structured, private peer assessment can meaningfully improve the equity of group-based assignment evaluation at Bishop Stuart University. Before the intervention, group assignments were evaluated through a uniform grading practice that did not reflect individual contribution and tolerated free-riding, consistent with the inequity Equity Theory predicts under such conditions. The intervention, implemented through an individually and privately administered peer rating tool with structured orientation, restored a measurable input-outcome balance: genuine, contribution-linked differentiation emerged in every one of the 55 groups studied, corroborated by participants' own reports of increased accountability and effort. The intervention nonetheless introduces its own challenges, reciprocal and majority-coalition marking, attempted collusion, and group-size effects, that must be actively managed rather than assumed away.

On this basis, the study recommends that Bishop Stuart University consider formalising structured peer assessment as institutional policy for group-based coursework, given the current absence of any documented guideline. Group sizes for assessed coursework should be kept close to ten members rather than fifteen to twenty, and formed deliberately according to seriousness and willingness to contribute rather than friendship. Orientation on the purpose and binding, final nature of peer ratings should be repeated at the start of every course unit that uses the tool, rather than assumed to transfer from one course to the next. Programme coordinators and lecturers should monitor for reciprocal and majority-coalition rating patterns and consider moderation mechanisms, such as flagging or adjusting statistically extreme outlier ratings within a group. Further research should track the same groups across multiple assessment cycles to establish whether the differentiation observed here persists, strengthens, or weakens with repeated use, and should extend this evaluative action research design to other faculties at Bishop Stuart University and to other Ugandan universities, to test how far these findings generalise beyond the Faculty of Education, Arts and Media Studies.

REFERENCES

 Copyright © The Author, All rights reserved

1. Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 2, pp. 267–299). Academic Press. https://doi.org/10.1016/S0065-2601(08)60108-2
2. Boud, D., & Falchikov, N. (Eds.). (2007). Rethinking assessment in higher education: Learning for the longer term. Routledge. https://doi.org/10.4324/9780203964309
3. Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (2022). Developing evaluative judgement in higher education: Assessment for knowing and producing quality work. Routledge. https://doi.org/10.4324/9781351115865
4. Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. SAGE Publications.
5. Cheng, W., & Warren, M. (2000). Making a difference: Using peers to assess individual students' contributions to a group project. Teaching in Higher Education, 5(2), 243–255. https://doi.org/10.1080/135625100114885
6. Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. Studies in Higher Education, 24(3), 331–350. https://doi.org/10.1080/03075079912331379935 39
7. Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. Educational Psychology Review, 32(2), 481–509. https://doi.org/10.1007/s10648-019-09510-3
8. Falchikov, N. (2005). Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education. Routledge.
9. Heng, T. T., Chu, S. K. W., & Kong, N. S. (2020). Peer assessment of individual contributions to a group project: Helping students develop self-regulation and evaluative judgement. Assessment & Evaluation in Higher Education, 45(3), 378–393. https://doi.org/10.1080/02602938.2019.1638996
10. Herr, K., & Anderson, G. L. (2022). The action research dissertation: A guide for students and faculty (3rd ed.). SAGE Publications.
11. Kaufman, D. B., Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. Journal of Engineering Education, 89(2), 133–140. https://doi.org/10.1002/j.2168-9830.2000.tb00507.x
12. Kemmis, S., McTaggart, R., & Retallic, J. (2004). The action research planner: Doing critical participatory action research (3rd ed.). SAGE Publications.
13. Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. Teaching in Higher Education, 11(3), 279–290. https://doi.org/10.1080/13562510600680582
14. Muzaale, T., Kasujja, J., & Kasekende, F. (2026). Rethinking Uganda's higher education landscape: Pedagogical innovation and technology-driven assessment. East African Journal of Arts and Social Sciences, 9(1), 235–253. https://doi.org/10.37284/eajass.9.1.4434
15. Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. (2004). Turning student groups into effective teams. Journal of Student Centered Learning, 2(1), 9–34.
16. Panadero, E., Jonsson, A., & Botella, J. (2019). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. Educational Research Review, 22, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004
17. Simms, A., & Nichols, T. (2014). Social loafing: A review of the literature. Journal of Management Policy and Practice, 15(1), 58–67.

 Copyright © The Author, All rights reserved

Volume 07, Issue 05 "September - October 2026"

ISSN 2583-0333

18. Topping, K. J. (2018). Using peer assessment to inspire reflection and learning. Routledge. https://doi.org/10.4324/9781315176857
19. Wen, M. L., & Tsai, C. C. (2006). University students' perceptions of and attitudes toward online peer assessment. Higher Education, 51(1), 27–44. https://doi.org/10.1007/s10734004-6375-8

 Copyright © The Author, All rights reserved