Research Article
Education and Learning

Group Discussion and Students' Academic Performance in History Among Secondary School Students in Mbeya City, Tanzania: A Descriptive Comparative Study

Hope John Sheyo

[The author informations are in the declarations section. This article is published by ETFLIN in Education and Learning, Volume 2, Issue 2, 2026, Page 67-79. DOI: 10.58920/edu0202779]

Received
11 August 2026

Revised
12 September 2026

Accepted
22 October 2026

Published
28 October 2026

Editor
Anik Ghufron

Keywords: Group discussion, teaching method, academic performance, History education, quasi-experimental design.

CC BY 4.0

AUTHOR (2026)

Abstract: This study compared History achievement between six secondary schools in Mbeya City, Tanzania: three using structured group discussion and three using the conventional question-and-answer method. Teaching method was assigned at the school level, and no school contributed students to both conditions, so teaching method is completely confounded with school identity; the study is therefore reported as descriptive, exploratory comparative evidence, not a statistical test of causal effectiveness. A quasi-experimental, pretest-posttest design was used with 120 students, balanced across Forms One to Three in each condition, over a six-to-seven-week intervention, with data from a 20-point History achievement test given before and after. Students in the group-discussion schools recorded a higher mean post-test score (M = 12.92, SD = 2.05) than students in the question-and-answer schools (M = 10.40, SD = 1.85), and a larger mean gain (7.73 versus 4.93 points). Student-level t-tests, ANCOVA, and Cohen's d are also reported (adjusted mean difference 2.48 points, $F(1, 117) = 47.86$; post-test d = 1.29), but, because treatment was assigned to only six school/class units and is fully confounded with school, these are descriptive summaries rather than confirmatory evidence of a treatment effect. Qualitative evidence from teacher interviews, focus groups, and observations was consistent with this pattern. The findings describe an observed difference between six participating schools under monitored conditions; they do not, alone, establish that group discussion caused it, since school-level factors could not be separated from teaching method in this design.

Introduction

Teaching methods play a significant role in shaping students' understanding, participation, and academic achievement. Education systems today increasingly favour learner-centred approaches that actively involve students in the learning process rather than leaving them to depend solely on teacher-centred methods. Among these approaches, group discussion has drawn attention in the literature as a strategy associated with interaction, knowledge sharing, critical thinking, and collaborative learning.

History as a subject asks learners to develop analytical abilities, interpretation skills, and the capacity to connect past events to contemporary society. Effective History teaching is more than the transmission of factual content; it is a process in which teachers and learners interact, reason, and reconstruct the past through evidence and interpretation. Yet the teaching of History in many secondary schools, in Tanzania and elsewhere in Africa, continues to face challenges tied to passive learning, where students mainly receive information from teachers with limited opportunities for active participation, an approach that weakens engagement and, in turn, understanding and performance.

Group discussion gives students room to exchange ideas, ask questions, explain concepts to peers, and build deeper understanding through social interaction. According to the sociocultural perspective of learning, knowledge is constructed through interaction between learners and their social environment, so involving students in collaborative activities such as group discussion may be associated with differences in academic outcomes.

In Tanzania, curriculum reforms have progressively shifted policy away from lecture-based instruction and rote memorisation toward learner-centred pedagogy that builds critical thinking, collaboration, and the application of knowledge (1). Despite this policy direction, empirical

studies continue to show that many teachers, History teachers included, still lean heavily on teacher-centred methods such as lecturing and simple question-and-answer techniques, partly because of limited pedagogical training, large class sizes, and inadequate instructional resources. This gap between policy intention and classroom practice is part of the motivation for comparing classroom outcomes under the two approaches, even where, as here, school-level factors cannot be fully separated from teaching method.

While learner-centred approaches have been studied fairly broadly in education generally, limited comparative evidence exists on group discussion specifically in relation to academic performance in History in Tanzanian secondary schools. Most existing studies focus on the general use of teaching methods, the availability of instructional resources, or the challenges teachers face implementing the competence-based curriculum, leaving classroom-level comparative evidence on group discussion and History achievement underexplored.

This study therefore compared History achievement in secondary schools in Mbeya City, Tanzania, between schools where structured group discussion was used and schools where the conventional question-and-answer method was used. Because the teaching method was assigned at the school level rather than to individual students, this comparison is descriptive and exploratory: it documents the pretest and post-test performance of students in the two sets of schools and the magnitude of the observed differences, without treating school as a randomised or otherwise separable factor from teaching method.

Theoretical Foundation

This study was informed principally by Vygotsky's Sociocultural Theory of Learning (2). The theory treats learning as a socially mediated process in which cognitive development is supported through interaction, language, and collaboration with others. Central to it is the Zone of Proximal Development (ZPD), the difference between what learners can accomplish independently and what they can accomplish with appropriate assistance from more capable peers or adults. Structured group discussion reflects this directly: learners explain ideas, question one another, encounter alternative interpretations, and jointly construct meaning. In History, such interaction can plausibly support understanding of relationships among events, causes, consequences, continuity, and change; the theory is used here to help interpret the observed classroom pattern, not to establish it as causally proven.

Bruner's discovery learning, Dewey's experiential education, Piaget's cognitive constructivism, and Freire's dialogic pedagogy are retained here only as complementary perspectives (3–9). They converge with the sociocultural perspective in emphasising active participation, dialogue, and learner engagement, but they were not treated as separate explanatory mechanisms in this study. Vygotsky's Sociocultural Theory therefore served as the principal interpretive lens for the observed pattern of association between structured peer interaction and students' measured History achievement, within the descriptive scope this study's design supports.

Literature Review

Group Discussion as a Teaching Method: Global Perspective

A substantial body of literature has examined group discussion across subjects and educational levels. Erskine and Tomkin (10), in one of the earliest experimental evaluations, found that group discussion within a complex teaching programme improved learners' comprehension compared with conventional instruction. More recent studies reach similar conclusions: Garside (11) compared lecture and group discussion strategies for developing critical thinking and found stronger gains among students exposed to discussion, while Slavin (12) argued that cooperative and group-based learning structures build responsibility for one's own learning and that of peers, associated with improved academic achievement across subjects.

Bonwell and Eison (13) further established that active learning strategies, small-group discussion included, increase student engagement and are associated with better retention of content than passive, lecture-based instruction. These findings have since been replicated across subject areas, including the sciences and social studies, where group discussion is consistently linked to improved test scores, greater engagement, and stronger communication skills, although some studies caution that these benefits depend on effective facilitation, time management, and equitable participation among group members.

Learner-Centred Pedagogy and Curriculum Reforms in Tanzania

Tanzania's adoption of the Competence-Based Curriculum (CBC) has placed strong policy emphasis on learner-centred, activity-oriented teaching. Nihuka (14) observed that although policy favours learner-centred approaches, teachers' understanding and practical application of methods such as group discussion remain limited due to inadequate training and resource constraints. Schweisfurth (7) cautions more broadly that learner-centred education, transplanted into developing-country contexts without adequate contextual adaptation, teacher preparation, and resource support, can shift from being a solution to becoming a practical problem for teachers who lack the capacity or classroom conditions to implement it effectively.

John (15), studying schools in Morogoro, Tanzania, found that the adoption of innovative, learner-centred teaching methods in secondary schools remains limited, with many teachers still relying on traditional lecture-based instruction despite curriculum reforms encouraging more participatory approaches. Findings such as these motivate classroom-level comparative evidence, of the kind this study attempts to provide, even though, as explained in the Methodology section, the present design cannot isolate the teaching method from the specific schools studied.

Teaching History and the Development of Critical Thinking

Literature on History education emphasises that the subject is inherently more than the transmission of factual content; it requires learners to analyse, interpret, and

reason about cause, change, and continuity. Studies in the Southern and East African region raise concerns that History teaching frequently leans on rote memorisation of facts over the development of critical thinking, which can make the subject appear dull to many learners. Ndomondo (16), studying History teachers in Tanzanian secondary schools, found that scarce instructional resources and heavy reliance on lecture-based transmission of historical narratives continue to constrain learners'historical thinking skills, including the ability to analyse, interpret, and communicate historical arguments.

Similarly, studies from Zimbabwe and Eswatini have found that History teachers rely predominantly on teacher-centred methods, textbooks as their sole resource, and assessment practices that reward recall over critical engagement with historical evidence. Harris and Reynolds (17) argue that meaningful history education must go beyond static factual recall and build learners' capacity to interpret and communicate historical thought, using second-order concepts such as causation and change. Bušljeta (18) adds that varied, well-structured instructional resources and interactive teaching strategies, discussion-based methods among them, help learners organise and internalise historical understanding more effectively than passive transmission of content.

Challenges Associated with Implementing Group Discussion

While the literature is broadly favourable toward group discussion, several studies caution that its benefits are not automatic and depend heavily on how the method is implemented. Time management has repeatedly emerged as a challenge, with studies noting that discussions can consume more classroom time than teacher-directed methods, particularly in the large classes typical of many African secondary schools. Unequal participation is another commonly cited concern, whereby more confident or academically stronger students dominate discussions while quieter students remain passive, a pattern that can undermine the collaborative learning the method is meant to produce unless teachers actively structure and monitor group interactions. Some studies further note that without clear guiding questions and defined roles, discussions can drift off-topic or fail to engage all members meaningfully.

These challenges suggest that the classroom value of group discussion is mediated by the quality of facilitation rather than being an inherent property of the method itself. Teachers who provide clear discussion prompts, rotate roles among group members, circulate to monitor group progress, and consolidate outcomes through whole-class synthesis are more likely to realise the cognitive and motivational benefits associated with the method. This matters particularly for History teaching, where discussions need to be anchored in specific historical evidence and guiding questions to avoid degenerating into unstructured, opinion-based exchanges that do not build genuine historical understanding.

Comparative Studies on Teaching Methods and Academic Performance

Comparative experimental and quasi-experimental studies across subjects provide further context. Several studies in Nigerian and Ghanaian secondary schools comparing group discussion, demonstration, and lecture methods across science subjects have found that group discussion and other activity-based methods are associated with higher measured academic achievement than lecture-based instruction, although the size of the association varies with subject matter, class size, and intervention duration. Studies in medical and higher-education settings similarly find that replacing passive instructional formats with small-group discussion is associated with improved knowledge outcomes, even though such settings differ considerably from secondary classrooms in learner maturity and class size.

Within Tanzania specifically, meta-analytic evidence synthesising multiple studies on teacher-centred *versus* learner-centred approaches across subjects reports a substantial average effect size favouring learner-centred methods, alongside significant variability across studies attributable to differences in design, subject area, and school context; many of the underlying primary studies share the same school-level assignment limitations discussed for the present study in the Unit of Assignment and the School-Treatment Confound section.

Recent Cooperative Learning Research in East Africa

Beyond the Tanzanian, History-specific literature reviewed above, recent quasi-experimental studies from the wider East African region offer further, more contemporary comparative evidence. Simesso (19), in a quasi-experimental study of secondary Chemistry students in Ethiopia, found that cooperative learning was associated with higher achievement and retention than conventional instruction, and, notably, addressed teacher-competence concerns by training two teachers of comparable qualification and experience for six days before the experimental teacher implemented the method. Kebede *et al*. (20) used a quasi-experimental jigsaw cooperative learning model in Ethiopian secondary Biology classrooms and reported meaningful achievement gains for the cooperative learning group over a conventionally taught control group. In a related tertiary-level study, Tadesse *et al*. (21) found that a cooperative learning pedagogical model was associated with enhanced engagement and learning outcomes among Ethiopian university students relative to conventional instruction, grounding their analysis in social interdependence and social constructivist theory.

These recent studies reinforce the theoretical and empirical expectation that structured, socially interactive teaching methods are associated with better outcomes than conventional instruction across subjects and levels within East Africa. However, almost all of this recent quasi-experimental evidence comes from Ethiopia and focuses on Chemistry, Biology, or general university coursework rather than on Tanzania or on History specifically, which sharpens the empirical gap motivating the present study. As with the present study, most of these designs assign treatment at the school or class level, and the corresponding identification limitations are not always made explicit.

Research Gap

Existing studies have generated useful insights on the

Table 1. Characteristics of participating schools and condition assignment.

School	Ownership	Gender composition	School size	Form/class	Condition
Idunda Secondary School	Government	Co-educational	Medium	Form One	Group discussion
UWATA Boys Secondary School	Private	Single-sex (boys)	Medium	Form Two	Group discussion
Samora Michael Secondary School	Government	Co-educational	Large	Form Three	Group discussion
Mbeya Day Secondary School	Government	Co-educational	Large	Form One	Question-and-answer
Uyole Secondary School	Government	Co-educational	Small	Form Two	Question-and-answer
Igawilo Secondary School	Private	Co-educational	Small	Form Three	Question-and-answer

Source: Field data, 2026

general adoption of learner-centred pedagogy in Tanzania, on the pedagogical challenges facing History teachers, and on cooperative and group-based learning in other subjects and countries within East Africa. Yet most of the Tanzania-and History-specific evidence remains descriptive or qualitative, relying on teachers'and students'perceptions rather than on measured, quantifiable classroom outcomes. Few studies have used a quasi-experimental pretest-posttest design, with attention to intervention fidelity and instrument psychometric properties, to directly compare the academic performance of students taught History through group discussion against those taught through conventional methods within the same subject and context. This study addresses that gap by providing systematic descriptive, comparative quantitative and qualitative evidence, strengthened by an explicit account of intervention fidelity and measurement quality, on group discussion and academic performance in History among secondary school students in Mbeya City, Tanzania, while explicitly acknowledging, rather than statistically resolving, the identification limits imposed by school-level assignment.

Methodology

Research Design

The study adopted a quasi-experimental design involving a non-equivalent comparison group, with a pretest-posttest structure. Random assignment of individual students was not feasible within existing school and classroom structures, so three schools were purposively assigned to the group-discussion condition and three to the question-and-answer condition (Population, Sample, and Sampling Procedure section). This non-equivalent group pretest-posttest structure is widely used in educational research for exactly this practical reason (22, 23), but it carries an important consequence for this study that is stated here explicitly rather than only in the limitations section: because condition was assigned at the school level, and no participating school contributed students to both conditions, teaching method is completely confounded with school identity in this design. Any observed difference between conditions could reflect the teaching method, other stable differences between the specific schools involved (for example resources, leadership, academic culture, or prior attainment patterns), or some combination of the two, and these cannot be statistically separated with six schools. The study is accordingly presented throughout as descriptive, exploratory comparative evidence characterising the six participating schools under the conditions described below, not as a design capable of isolating a population-level causal effect of group discussion. The Unit of Assignment and the School-Treatment Confound section sets out this limitation, and a related structural confound involving Form and topic, in full.

The pretest was used to characterise the two sets of schools before the intervention; a nonsignificant student-level baseline comparison is reported in the Results and Findings section and is interpreted only as indicating that no statistically significant pretest difference was detected at the student level, not as proof of equivalence between the schools themselves. The quantitative component was complemented by qualitative data from teacher interviews and student focus group discussions, used to explain and contextualise the descriptive quantitative pattern.

Study Area

The study was conducted in secondary schools within Mbeya City Council, Mbeya Region, Tanzania. Mbeya City was chosen for the diversity of its secondary schools, which span public and private institutions across low-, average-, and high-performing categories in History, and operate under different administrative structures and resource levels. This diversity offered a realistic setting for comparing group discussion and question-and-answer teaching under ordinary classroom conditions, and a meaningful basis for considering how school-level factors, such as ownership and resource availability, relate to the implementation of learner-centred methods. Mbeya City was therefore a strategic choice for gathering relevant, reliable, and comprehensive comparative data, understood within the limits described in the Research Design section.

Population, Sample, and Sampling Procedure

The target population comprised History teachers and secondary school students in six purposively selected secondary schools within Mbeya City Council: Idunda, Igawilo, Samora Michael, UWATA Boys, Mbeya Day, and Uyole Secondary Schools. These schools were selected to represent variation in location (urban and peri-urban), ownership (public and private), gender composition (single-sex and co-educational), student population size, and socio-economic background, ensuring broad representation of the different educational contexts within the city.

For the comparative component reported here, all six schools took part, with one sampled group of students selected from each; three schools (Idunda, UWATA Boys, Samora Michael) were assigned to the group-discussion condition and three (Mbeya Day, Uyole, Igawilo) to the question-and-answer condition. The sampled groups represented Forms One, Two, and Three in both

conditions, giving 60 students per condition and roughly 20 students per sampled group. Assignment of the six schools to condition was purposive rather than random: the researcher considered academic comparability and prior History performance across the two sets of three schools, while preserving contextual variation in ownership, gender composition, and school size within each condition. No formal randomisation or matching algorithm was used, and no independent baseline school-level comparability statistics (for example, prior school-wide History pass rates) beyond the researcher's qualitative judgement are available to substantiate the allocation; this is acknowledged explicitly as a limitation in the Study Limitations section. **Table 1** summarises the characteristics of the six participating schools and their condition assignment.

The sampled groups were not the entire student population of each school's relevant Form. In several participating schools, a single Form enrolled well over 120 students in total; from each school, the researcher drew a sample of approximately 20 students from the relevant Form's total population, rather than using an entire, pre-existing class section. Once selected, each sampled group was taught together by the same visiting teacher for the full six-to-seven-week intervention, functioning as a single instructional unit for the duration of the study. Because the sampled students were not necessarily an intact, long-standing class section prior to the study, the degree of pre-existing peer familiarity available to a naturally intact class cannot be assumed; this is noted as a limitation in the Study Limitations section.

Table 1 shows that the two conditions were distributed across three schools each, with Forms One, Two, and Three represented equally in both. It also shows, directly, the central limitation of this design: the group-discussion condition and the question-and-answer condition consist of two entirely separate sets of three schools, with no school contributing students to both. Teaching method and school identity are therefore completely confounded, and this structure, together with the related Form/topic confound, is set out in full in the Unit of Assignment and the School-Treatment Confound section and carried through the interpretation of every result reported in Results and Findings section and Discussion of Findings section .

Teacher and Control of Teacher Effects

To reduce the possibility that differences between conditions reflected differences in teacher competence rather than teaching method, the same experienced visiting History teacher delivered instruction in all six participating schools, teaching the three group-discussion schools using structured group discussion and the three question-and-answer schools using the conventional method. Using one teacher across both conditions reduced between-teacher variation in qualification, experience, personality, and general instructional competence, though it does not address the school-level confound described in the Unit of Assignment and the School-Treatment Confound section, since school identity itself, not only the teacher, differs systematically between conditions. The researcher did not personally teach any of the six schools; the researcher coordinated the study, developed the achievement test and lesson materials for both conditions, oriented the visiting teacher to the group-discussion protocol described in Intervention Description and Fidelity section, and conducted classroom observations to monitor implementation. The teacher worked alongside the regular History teachers in the participating schools, who retained routine responsibility for their classes.

Using a single teacher across both conditions also raises a distinct methodological concern: the teacher necessarily knew which schools were being taught by which method, creating a risk of teacher- or experimenter-expectancy effects, where the teacher might unconsciously invest more energy or enthusiasm in the group-discussion lessons, and a risk of treatment diffusion, where techniques characteristic of group discussion might inadvertently carry over into question-and-answer lessons, or *vice versa*. To mitigate this, the teacher followed detailed, written lesson plans specific to each condition, prepared in advance by the researcher, and the classroom observation checklist described in Intervention Description and Fidelity section explicitly checked question-and-answer lessons for any unintended use of small-group structuring or discussion-based techniques, alongside checking group-discussion lessons for protocol adherence; no instances of cross-condition contamination were recorded during the observed lessons. This concern is revisited in the limitations section.

Research Instruments

An achievement test in History, developed by the researcher, served as the main quantitative instrument, administered as a pretest before the intervention and a post-test after it. The test comprised 20 items, combining multiple-choice and short structured-response questions built around the History topics covered during the intervention, designed to assess both factual recall and higher-order historical thinking, such as explaining the causes and consequences of historical events. The pretest and post-test were constructed as parallel forms covering the same topics and cognitive levels, to limit direct item-memorisation effects between the two testing occasions. Focus group discussions and a semi-structured interview with the History teacher supplied complementary qualitative insight that helped explain and contextualise the quantitative findings.

Intervention Description and Fidelity

The intervention ran over six to seven weeks and was designed around selected History topics from the Tanzania History syllabus (24): for Form One, socio-political organisation and education in pre-colonial African societies; for Form Two, the origin, development, and effects of the slave trade from the fifteenth to the nineteenth century; and for Form Three, colonial social services. Because each Form follows a different official syllabus topic, and each Form is represented by exactly one group-discussion school and one question-and-answer school (**Table 1**), school, Form, and topic are also confounded with one another within each Form-level comparison; this structural point is developed in the Unit of Assignment and the School-Treatment Confound section. The study followed the syllabus allocation of 14

periods for the Form One topic, 13 periods for the Form Two topic, and 12 periods for the Form Three topic. Since the syllabus defines one period as 40 minutes, the intervention used 80-minute instructional sessions, equivalent to two periods, where the timetable allowed. This gave roughly seven weeks for Form One, six and a half weeks for Form Two, and six weeks for Form Three, all within the planned six-to-seven-week window.

The same trained, experienced visiting teacher, who was not the researcher, delivered the assigned instructional method in all six schools according to a pre-arranged weekly timetable. Because all six schools were located within Mbeya City, movement between them was feasible: each school received one visit per week, School A on Monday through School F on Saturday, with School F taught between 10:00 a.m. and 11:30 a.m. on Saturdays. In the three group-discussion schools, students worked in structured mixed-ability groups with defined roles, guiding questions, peer explanation, group presentation, and whole-class consolidation; in the three question-and-answer schools, the same teacher used the conventional approach. Written lesson plans and fidelity observations helped maintain consistency across the six schools.

Before the intervention, the visiting teacher was oriented by the researcher to the study procedures and the group-discussion protocol, covering formation and management of mixed-ability groups, questioning and probing, time management, presentation of group findings, and whole-class consolidation, along with guidance on maintaining a clear distinction between the group-discussion protocol used in one set of schools and the conventional question-and-answer method used in the other. The researcher monitored implementation through structured classroom observations. The weekly rotation across the six schools let the same teacher teach one participating school per scheduled day rather than all six on the same day, and, since all schools were within Mbeya City, travel demands between them remained manageable.

Validity and Reliability

Content validity of the comprehensive history achievement test was rigorously established through expert review by two experienced History educators. They systematically evaluated how well individual test items aligned with the intervention syllabus content and the specific Bloom's taxonomy levels targeted by the study, ensuring that their constructive feedback directly informed minor revisions to item wording before piloting. A formal pilot test followed with 30 Form Two students from a secondary school not involved in the main study, who were comparable in academic level to the main sample. Pilot data were used to compute item difficulty indices, which ranged from 0.32 to 0.78 (mean = 0.55), indicating an acceptable spread of easy, moderate, and more challenging items, and item discrimination indices, of which 85 percent of items exceeded the conventional acceptability threshold of 0.30, with the remaining items revised or replaced before main data collection. Internal consistency reliability of the achievement test, computed using Cronbach's alpha on the pilot data, was .82, indicating good reliability for a researcher-developed classroom achievement test of this length.

Data Collection Procedure

Before the intervention, students in both sets of schools sat the same 20-point pretest to characterise their starting academic performance in History. The group-discussion schools were then taught the selected Form-specific syllabus topic using structured group discussion, following the protocol in Intervention Description and Fidelity section, while the question-and-answer schools were taught the corresponding topic using the conventional method. The intervention followed the period allocations planned for each Form and was delivered in 80-minute sessions, equivalent to two 40-minute periods, within the six-to-seven-week schedule, with each participating school receiving one scheduled teaching visit per week. After the topic-specific intervention was complete, both sets of schools sat the post-test, covering the same content domain and cognitive levels as the pretest. Qualitative data collection, described in Qualitative Data Collection and Analysis section, followed immediately after.

Qualitative Data Collection and Analysis

To complement and help explain the quantitative results, qualitative data were gathered through an in-depth semi-structured interview with the visiting History teacher, uniquely placed to compare the delivery and reception of both methods across all six schools, and six focus group discussions, one per participating school, three from group-discussion schools and three from question-and-answer schools, each with seven to eight students selected through purposive, maximum-variation sampling to include high, average, and lower-performing students based on pretest scores. Interviews and focus group discussions were conducted after the post-test, audio-recorded with participants' consent, transcribed verbatim, and analysed thematically following the six-phase approach of Braun and Clarke (25): familiarisation with the data, generation of initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the final analytic narrative. Two researchers independently coded a subset of transcripts and reconciled coding discrepancies through discussion, strengthening the trustworthiness of the qualitative analysis.

Consistent with the descriptive, exploratory framing of this study, qualitative evidence was used primarily to explain and contextualise the quantitative pattern rather than to constitute independent confirmatory evidence of a treatment effect. Three themes proved especially relevant: peer explanation and clarification, active participation and confidence, and the importance of teacher facilitation and classroom conditions. Representative quotations from the interview and focus group data are integrated into the results narrative to show how participants explained the classroom processes they observed; these qualitative accounts describe the group-discussion schools and cannot, on their own, rule out school-level explanations for the quantitative pattern either.

Data Analysis Procedure

Quantitative data from the pretest and post-test were summarised using means and standard deviations, and, for descriptive completeness, independent-samples *t*-tests, paired-samples *t*-tests, one-way ANCOVA with pretest

score as covariate, and Cohen's *d* (26) were also computed at the student level. These student-level statistics are reported in Results and Findings section, but they are not treated as confirmatory tests of a population-level teaching-method effect. Two design features prevent that stronger interpretation and are set out in full in the Unit of Assignment and the School-Treatment Confound section: first, treatment was assigned to only six school/class units (three group-discussion, three question-and-answer), so the 120 student-level observations represent six, not 120, independent treatment assignments; second, the group-discussion and question-and-answer conditions consist of entirely separate sets of three schools each, so teaching method is completely confounded with school identity, and no statistical adjustment available with six schools can separate the two. Given this, a conventional multilevel or cluster-robust model was not fitted as a corrective analysis: six clusters are too few to estimate such a model with acceptable precision, and in any case no such model can resolve a complete confound between treatment and school. The *t*-tests, ANCOVA, *p*-values, and Cohen's *d* reported in Results and Findings section are accordingly presented as descriptive, exploratory summaries of the observed student-level data, not as inferential evidence establishing the effectiveness of group discussion.

Unit of Assignment and the School-Treatment Confound

Two related but distinct design limitations govern how the quantitative results in Results and Findings section should be read, and they are stated here explicitly rather than left as an implicit caveat.

The first is a statistical, unit-of-assignment problem. Treatment was assigned to six school/class units, not to 120 individual students. Students who share a school are not independent of one another for the purpose of estimating a teaching-method effect: they share a teacher, timetable, school environment, leadership, and academic culture, among other factors. This means the effective number of independent treatment assignments in this study is six, not 120, and the student-level *p*-values, confidence intervals, and standard errors reported in Results and Findings section understate the true uncertainty of the comparison. With only three units per condition, a multilevel or cluster-robust re-analysis would itself be too imprecise to serve as a corrective solution, so none is presented; the limitation is instead addressed by declining to treat the student-level inferential statistics as confirmatory evidence.

The second, and more fundamental, limitation is that teaching method is completely confounded with school in this design. The three group-discussion schools (Idunda, UWATA Boys, Samora Michael) and the three question-and-answer schools (Mbeya Day, Uyole, Igawilo) are entirely distinct sets; no school contributed students to both conditions (**Table 1**). Consequently, any observed difference between conditions could reflect the teaching method, or it could reflect other stable differences between these specific schools, such as resources, leadership, academic culture, attendance patterns, or prior attainment, and these explanations cannot be statistically separated with six schools. A related structural confound arises from the History syllabus itself: each Form was taught a different topic (Intervention Description and Fidelity section), and each Form is represented by exactly one group-discussion school and one question-and-answer school, so school, Form, and topic are also confounded with one another within each Form-level comparison.

Together, these two limitations mean that the quantitative comparisons in this manuscript cannot support a causal claim that group discussion produced the observed differences. They are reported and interpreted throughout, in sections Results and Findings and Discussion of Findings, as descriptive, exploratory comparative evidence characterising the six participating schools under the specific conditions of this study, not as a statistically validated test of the teaching method's effectiveness.

Ethical Considerations

Administrative permission to conduct the research was obtained from the relevant regional administrative authority before data collection, authorising research activities in the selected secondary schools. Formal ethical clearance for the study was also obtained prior to data collection; in keeping with this journal's editorial policy, the approving body and clearance reference are disclosed separately through the manuscript submission portal rather than within the manuscript text. Informed consent was obtained from participating teachers and students, and, for students under 18 years, from their parents or guardians as well. Participants were informed of the study's purpose and procedures, the voluntary nature of participation, and their right to withdraw without negative consequences. Confidentiality and anonymity were maintained by coding participant information and reporting findings in aggregate form.

The researcher also took deliberate steps to avoid disadvantaging the question-and-answer schools, ensuring that the conventional method was delivered by the same competent, experienced teacher and in line with normal classroom practice, rather than being deliberately weakened for comparison. After data collection was complete, the researcher shared a summary of the group discussion protocol and lesson materials with the participating schools, so that the study's practical value extended beyond the group-discussion condition.

Limitations of the Study

This study has several important limitations, and the first two are the most consequential for how its findings should be interpreted.

First, and most seriously, teaching method is completely confounded with school: the three group-discussion schools and three question-and-answer schools are entirely distinct sets, and the design cannot statistically separate a genuine teaching-method effect from other stable differences between these specific schools (Unit of Assignment and the School-Treatment Confound section). This is a structural limitation of the design itself, not merely a matter of interpretive caution, and it is the main reason this study is presented as descriptive, exploratory comparative evidence rather than as a causal test of effectiveness.

Second, and related, treatment was assigned to only

six school/class units, so the 120 student-level observations reflect six, not 120, independent treatment assignments; conventional student-level *p*-values, confidence intervals, and effect sizes understate the true uncertainty of the comparison, and a multilevel or cluster-robust re-analysis is not feasible with only three units per condition.

Third, each Form was taught a different History topic and is represented by only one group-discussion school and one question-and-answer school, so school, Form, and topic are also confounded with one another within each Form-level comparison; this compounds the school-treatment confound described above.

Fourth, allocation of schools to condition was purposive rather than random: the researcher considered academic comparability, prior History performance, and contextual variation in ownership, gender composition, and school size when assigning the three group-discussion and three question-and-answer schools, but no formal randomisation or matching algorithm was used, and no independent baseline school-level comparability statistics are available beyond this qualitative judgement.

Fifth, using one experienced visiting teacher across both conditions reduced teacher-level variation in qualification and experience but did not eliminate the possibility of teacher-expectancy effects, since the teacher knew the assigned method for each school; it also does not address the school-level confound itself, since school identity, not only the teacher, differs systematically between conditions.

Sixth, the six-to-seven-week intervention was relatively short and closely monitored, so the descriptive differences reported here may partly reflect novelty, heightened engagement, or test-retest sensitisation, since the same general test format was used at pretest and post-test, rather than a stable, durable difference attributable to teaching method.

Finally, the sampled groups in each school were drawn from a larger Form population rather than constituting an intact, pre-existing class section, so the same degree of pre-existing peer familiarity available to a naturally intact class cannot be assumed; and the achievement test measured the specific learning outcomes represented by the 20-point assessment, not a separately validated measure of critical thinking or long-term retention.

Taken together, these limitations mean the study should be read as descriptive, exploratory evidence about six specific schools under closely monitored conditions, not as a validated test of the causal effectiveness of group discussion as a teaching method. Replication with a substantially larger number of schools, ideally with several schools per condition per Form and a reproducible, and where feasible randomised, allocation procedure, would be needed to separate teaching-method effects from school-level effects and to support formal multilevel or cluster-robust inference.

Results and Findings

This section presents the quantitative comparison between the group-discussion schools and the question-and-answer schools, and the qualitative evidence used to contextualise the observed pattern. Consistent with section Data Analysis Procedure and Unit of Assignment and the School-Treatment Confound, all statistics reported here are descriptive summaries of the observed student-level data; because teaching method is completely confounded with school, and treatment was assigned to only six school/class units, none of the statistics below are presented as confirmatory evidence of a causal effect of group discussion. The achievement test had a maximum possible score of 20 points.

These findings respond to the study's objective of comparing History achievement between the group-discussion schools and the question-and-answer schools in Mbeya City. Academic performance was assessed using pretest and post-test scores for both sets of schools.

Descriptive and Between-Group Comparisons

Table 2 consolidates the descriptive and between-group comparisons; scores are reported on the 20-point achievement test. At the student level, the pretest comparison detected no statistically significant baseline difference between the two sets of schools, while the post-test comparison showed a higher mean in the group-discussion schools. Because of the school-treatment confound described in the Data Analysis Procedure section, these comparisons are reported descriptively.

Students in the question-and-answer schools improved from M = 5.47 to M = 10.40, a gain of 4.93 points, while students in the group-discussion schools improved from M = 5.18 to M = 12.92, a gain of 7.73

Table 2. Descriptive statistics and between-group comparisons for pretest and post-test scores.

Comparison	Group	N	Mean (SD)	Mean gain	t	df	p	95% CI
Pretest	Question-and-answer	60	5.47 (2.57)	—	0.61	118	.540	−1.20 to 0.62
Pretest	Group discussion	60	5.18 (2.48)	—				
Post-test	Question-and-answer	60	10.40 (1.85)	—	7.05	118	<.001	1.81 to 3.23
Post-test	Group discussion	60	12.92 (2.05)	—				
Within-group change	Question-and-answer	60	5.47 → 10.40	4.93 (t = 11.57, p <.001)	—	59	<.001	—
Within-group change	Group discussion	60	5.18 → 12.92	7.73 (t = 18.73, p <.001)	—	59	<.001	—

Source: Field data, 2026. **Note:** all statistics are descriptive/exploratory; treatment is completely confounded with school. 95% CIs are approximate, based on the reported group means and standard deviations.

points. The raw post-test difference was 2.52 points (95% CI 1.81 to 3.23), favouring the group-discussion schools, while the pretest difference was small and not statistically significant at the student level (95% CI −1.20 to 0.62). In short, the group-discussion schools recorded higher measured achievement at post-test than the question-and-answer schools; because of the school-treatment confound set out in the Data Analysis Procedure section, this comparison describes an observed difference between the six participating schools and is not presented as confirmed evidence of a teaching-method effect.

ANCOVA and Standardised Differences (Descriptive)

For descriptive completeness, ANCOVA was used to compare post-test scores between groups after adjusting for pretest score at the student level, as explicitly presented in **Table 3**. The homogeneity-of-regression-slopes assumption was checked and held before the model was fitted; inspection of residuals did not indicate a substantial departure from the model's assumptions, and no influential observation was identified as grounds for excluding a participant from this comprehensive statistical analysis. The adjusted post-test means were 10.42 for the question-and-answer schools and 12.90 for the group-discussion schools, clearly demonstrating the observed group differences.

At the student level, the ANCOVA-adjusted comparison showed a higher adjusted mean in the group-discussion schools than in the question-and-answer schools, $F(1, 117) = 47.86$, $p < .001$, adjusted mean difference 2.48 points. The post-test standardised difference ($d = 1.29$) and mean-gain standardised difference ($d = 0.86$) are both numerically large by conventional statistical benchmarks. However, because treatment was assigned to only six school/class units and is completely confounded with school identity (Data Analysis Procedure section), these figures are reported here as unadjusted, student-level descriptive summaries of the observed data, not as evidence of the magnitude of a causal teaching-method effect; conventional labels such as "large effect" in the inferential sense are deliberately not applied to them.

These statistics should be read together with Data Analysis Procedure section. The six sampled groups were nested within six schools, three per condition, so the effective number of independent treatment assignments is six, not 120: intra-class correlation among students within the same school means the student-level *p*-values, confidence intervals, and *d* estimates overstate the precision of the comparison. Separately, and more fundamentally, because the group-discussion and question-and-answer schools are entirely distinct sets of schools, no statistical adjustment available with six schools can separate a genuine teaching-method effect from other stable differences between these specific schools. Both limitations apply simultaneously; correcting only the first (for example, by attempting a cluster-robust adjustment) would not resolve the second. The descriptive pattern, students in the group-discussion schools scoring higher, is consistent across the raw post-test comparison, the paired-samples comparisons, and the ANCOVA-adjusted comparison, but its cause cannot be attributed to the teaching method with this design.

Summary of Findings

The descriptive findings show improvement in both sets of schools from pretest to post-test, with a larger mean gain in the group-discussion schools (7.73 *versus* 4.93 points). At post-test, the group-discussion schools had a higher mean score, and the student-level ANCOVA showed a numerically large adjusted difference. No statistically significant student-level pretest difference was detected between the two sets of schools. Qualitative evidence helped contextualise this descriptive pattern, pointing to peer explanation, active participation, and structured teacher facilitation as features the teacher and students associated with the group-discussion lessons.

Because teaching method is completely confounded with school, and treatment was assigned to only six school/class units, these findings are summarised as an observed comparative pattern across the six participating schools under closely monitored conditions, not as statistically confirmed evidence that group discussion caused the difference (Unit of Assignment and the School-Treatment Confound section).

Table 3. Student-level ANCOVA and standardised differences for post-test achievement (descriptive).

Analysis / effect	Estimate	F	df	p	Partial η^2 / d	95% CI
ANCOVA: pretest covariate	—	38.42	1, 117	<.001	.247	—
ANCOVA: group term	Adjusted mean difference = 2.48	47.86	1, 117	<.001	.290	1.77 to 3.19 (approx.)
Adjusted mean: question-and-answer	10.42	—	—	—	—	—
Adjusted mean: group discussion	12.90	—	—	—	—	—
Post-test standardised difference (d)	1.29	—	—	—	descriptive only	0.89 to 1.68 (approx.)
Mean-gain standardised difference (d)	0.86	—	—	—	descriptive only	—

Source: Field data, 2026. **Note:** these are unadjusted, student-level descriptive summary statistics, not confirmatory inferential estimates of a teaching-method effect; Cohen's d (26) magnitude labels (small/medium/large) are omitted here for that reason.

Discussion of Findings

Students in the group-discussion schools recorded higher post-test scores than students in the question-and-answer schools. Both sets of schools improved over the intervention period, but the group-discussion schools recorded the larger mean gain (7.73 points *versus* 4.93 points). The raw post-test difference was 2.52 points, and the ANCOVA-adjusted, student-level difference was approximately 2.48 points; the consistency of these descriptive estimates shows a stable comparative pattern across the six participating schools. Because the design assigns teaching method at the school level, and no school contributed students to both conditions, this pattern is discussed here as an observed association between the group-discussion schools and higher measured History achievement, not as evidence that group discussion itself caused the difference.

A central alternative explanation cannot be ruled out with this design: the three group-discussion schools and three question-and-answer schools may differ from one another in ways unrelated to teaching method, for example in resources, school leadership, academic culture, attendance patterns, or the prior attainment composition of their students, and any of these could plausibly contribute to the observed difference. The study's data cannot distinguish this explanation from a genuine teaching-method effect, because school and teaching method are completely confounded (Unit of Assignment and the School-Treatment Confound section). The qualitative evidence below describes what the teacher and students in the group-discussion schools reported experiencing, which is consistent with a facilitation-related explanation, but it does not, on its own, rule out school-level explanations either.

Read descriptively, this pattern is consistent with Vygotsky's Sociocultural Theory (2), which positions learning as socially mediated through interaction, language, and guided participation. In this intervention, students in the group-discussion schools were required to articulate ideas, respond to peers, and negotiate interpretations of historical content. The qualitative evidence told a similar story. One student explained: "I like group discussion because it helps me learn from my fellow students and understand topics that I did not understand during the lesson." Another noted that "group discussion makes learning History more interesting because everyone can share different ideas and examples." These accounts illustrate the peer-explanation and interaction processes the sociocultural perspective emphasises, and they are consistent with, though they do not independently establish, the quantitative pattern of higher measured achievement in the group-discussion schools.

These descriptive results broadly resemble prior empirical literature on discussion and active learning. Garside (11) reported advantages of group discussion over lecturing for critical-thinking development, while Slavin (12) and Bonwell and Eison (13) described benefits associated with cooperative and active learning. Erskine and Tomkin (10) similarly reported improved comprehension under group discussion. These studies offer contextual resemblance to the present descriptive pattern, though differences in subject, learner population, outcome measures, and, in many cases, similar school-level assignment limitations to those affecting this study mean their findings should not be treated as independent confirmation of a causal effect here.

Within the Tanzanian context, these descriptive findings speak to the learner-centred orientation of curriculum reforms. Nihuka (14) and John (15) reported continuing challenges implementing learner-centred methods, tied to teacher preparation and resource limitations. Ndomondo (16) likewise identified resource constraints and transmission-oriented History teaching as important contextual issues. In this study, the teacher interview and classroom observations indicated that the question-and-answer approach supported recall but offered fewer opportunities for extended peer interaction; this is offered as a plausible, teacher-reported explanation for part of the observed pattern, alongside the school-level explanations that cannot be excluded.

These findings also carry relevance for how History education is discussed, since meaningful historical learning involves more than recalling isolated facts. Harris and Reynolds (17) and Bušljeta (18) emphasise interpretation, communication, and the effective organisation of historical learning resources. The present achievement test included items requiring understanding beyond simple recall, but the analysis reported a total achievement score rather than a separate validated measure of critical thinking. The higher scores in the group-discussion schools are therefore read as higher measured History achievement in those schools, not as evidence that group discussion independently improved critical-thinking ability.

Design Limits on Causal Interpretation

Two design features are central to how the observed group difference should be read, and both were introduced in Unit of Assignment and the School-Treatment Confound section. First, treatment was assigned to six school/class units, so the 120 student-level observations represent six, not 120, independent treatment assignments; the student-level inferential statistics in the Results and Findings section accordingly overstate the precision of the comparison. Second, and more fundamentally, the three group-discussion schools and three question-and-answer schools are entirely separate sets of schools: teaching method is completely confounded with school identity, and this study's data cannot statistically separate a genuine effect of group discussion from other stable differences between these specific schools. A related structural confound compounds this: each Form was taught a different History topic, and each Form is represented by exactly one group-discussion school and one question-and-answer school, so school, Form, and topic are also confounded with one another. Taken together, these features mean the study documents an observed comparative pattern across six specific schools under closely monitored conditions; it does not, and by design cannot, establish that group discussion caused this pattern rather than some combination of the schools' other characteristics.

Theoretical Implications

The theoretical implication sits most directly with Vygotsky's Sociocultural Theory (2). The observed pattern, higher measured History achievement in the

group-discussion schools, is consistent with the proposition that learning can be supported through social interaction, dialogue, and guided participation within learners'Zones of Proximal Development, and the qualitative accounts of peer explanation and idea sharing provide contextual, descriptive support for these processes within the group-discussion schools. Given the school-treatment confound described above, this should be read as a finding consistent with the theory rather than a test of it: the design cannot rule out that some or all of the observed difference reflects factors unrelated to social interaction in the classroom. Bruner, Dewey, Piaget, and Freire remain useful complementary perspectives, but this study did not test distinct mechanisms derived from each theory, so no separate empirical claims are made for them.

The findings also support a contextual reading of learner-centred pedagogy. Schweisfurth (7) cautions that learner-centred approaches can become difficult to implement when classroom conditions and teacher preparation are insufficient, and the present qualitative findings similarly indicate that group discussion's perceived usefulness depended on teacher guidance, manageable group organisation, adequate time, and opportunities for balanced participation, as reported by the teacher and students in the group-discussion schools.

Interpreting the Magnitude of the Descriptive Differences

The unadjusted, student-level standardised differences reported in ANCOVA and Standardised Differences (Descriptive) section (d = 1.29 for the post-test comparison, d = 0.86 for the mean-gain comparison) are numerically large by conventional statistical benchmarks, but, consistent with Data Analysis Procedure section, they are not presented here as evidence of the magnitude of a causal intervention effect. They are calculated from student-level observations despite treatment being assigned to six school/class units, so they inherit both the pseudoreplication problem described in Unit of Assignment and the School-Treatment Confound and the complete confounding of treatment with school; either issue alone would be sufficient reason not to interpret them as confirmed effect sizes in the ordinary inferential sense.

Plausible contributors to the magnitude of the descriptive difference include genuine classroom processes associated with group discussion, as well as factors unrelated to teaching method: notable differences between the specific schools involved, the short, closely monitored six-to-seven-week intervention period, possible novelty or teacher-expectancy effects, and the parallel pretest/post-test test format employed here. Recent East African studies, including Simesso (19), Kebede *et al.* (20), and Tadesse *et al.* (21), report broadly similar descriptive patterns favouring cooperative and discussion-based learning, which highlight pedagogical benefits; but they differ from this study in subject, educational level, and context, and many share comparable school- or class-level assignment limitations; consequently, they are offered here primarily as contextual resemblance, not as independent statistical confirmation of the magnitude observed in this empirical study.

Practical and Policy Implications

Given the design limitations set out above, the practical implications of this study are offered tentatively rather than as settled guidance. Teachers may consider structured group discussion as one instructional approach within History teaching, particularly where classroom conditions permit effective facilitation, clear guiding questions, defined group roles, and whole-class consolidation, since the descriptive pattern in this study and in the broader literature is consistent with such benefits. However, because this study's design cannot separate the teaching method from the specific schools studied, it does not by itself justify a policy recommendation that group discussion be adopted on effectiveness grounds. Further studies using a larger number of schools, ideally several schools per condition per Form, and allocation procedures that permit separation of school and treatment effects, are needed before stronger causal conclusions, and correspondingly stronger policy recommendations, can be drawn.

For curriculum developers, including the Tanzania Institute of Education, this study's descriptive pattern is consistent with, and may usefully inform, continued attention to learner-centred, activity-oriented approaches within the History syllabus, for example through discussion-based lesson exemplars and assessment approaches that reward historical reasoning; such steps are supported by the wider literature reviewed in the Literature Review section, and this study adds a Tanzania- and History-specific descriptive data point to that literature, without on its own establishing the causal case for any single instructional method.

Conclusion

This study compared History achievement between six secondary schools in Mbeya City, Tanzania, three using structured group discussion and three using the conventional question-and-answer method. Under the monitored conditions of this study, students in the group-discussion schools recorded higher post-test scores and larger pretest-to-post-test gains than students in the comparison schools. However, because teaching method was assigned at the school level and is completely confounded with school identity, and because Form and topic characteristics could not be separated from school within this six-school design, these differences cannot be interpreted as an isolated causal effect of group discussion. The findings are best read as descriptive, exploratory comparative evidence from six secondary schools in Mbeya City, Tanzania, consistent with, but not establishing, the broader hypothesis that structured group discussion supports History achievement. Replication with a substantially larger and more balanced set of schools, ideally allowing several schools per condition per Form and a reproducible, and where feasible randomised, allocation procedure, would be needed before stronger causal claims can be drawn.

Recommendations

Based on the descriptive findings and conclusion of this study, and in view of its design limitations, the following recommendations are made.

First, teachers may consider structured group

discussion as one instructional approach within History teaching, particularly where classroom conditions permit effective facilitation, clear guiding questions, defined group roles, and whole-class consolidation; this is offered as a reasonable, evidence-consistent option for practice, not as a proven prescription.

Second, school administrations and education authorities may support discussion-based, learner-centred teaching by providing adequate instructional time, manageable group organisation, and relevant teaching and learning resources, where they choose to adopt such approaches for reasons consistent with the broader literature reviewed in Literature Review section.

Third, teacher training institutions and continuous professional development programmes should strengthen teachers' skills in designing, facilitating, and assessing group discussion activities, so that participation stays balanced and learning tasks remain focused, independently of the causal question this study cannot resolve.

Fourth, curriculum developers, including the Tanzania Institute of Education, may consider practical exemplars showing how structured discussion can be aligned with specific History topics and assessed learning outcomes, informed by the accumulating descriptive and experimental literature rather than by this single study alone.

Finally, and most importantly for the causal question this study set out to explore, further studies using a substantially larger number of schools, ideally with several schools per condition per Form, allocation procedures that permit separation of school and treatment effects, and, where feasible, randomised assignment at the school or class level, are needed before stronger causal conclusions are drawn. Such studies should also report baseline school-level comparability evidence, consider delayed post-tests to assess retention, and examine how group size, facilitation style, and student characteristics relate to classroom outcomes under group discussion.

Declaration

Author Information

Hope John Sheyo *(*Corresponding author)*, Teofilo Kisanji University, Faculty of Education, PO box 1104, Mbeya, Tanzania. **Contribution:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft, Visualization, Validation, Supervision, Software, Resources.

Acknowledgment

During the preparation of this manuscript, the author used Claude (Anthropic), an AI-based large language model, to support specific aspects of the writing and editing process. Claude was used to assist with structuring and organizing content in accordance with journal formatting requirements, improving clarity and grammatical accuracy of the English text, and verifying the accuracy and traceability of cited references. The tool was also used to support the organization of statistical results (e.g., ANCOVA output and effect size reporting) into narrative form based on data and analyses generated and interpreted by the author. All research design, data collection, statistical analysis, interpretation of findings, and scholarly conclusions are the sole intellectual work of the author. Claude was not used to generate original research findings, data, or citations without verification. All AI-assisted content was critically reviewed, edited, and verified by the author prior to submission, and the author take full responsibility for the accuracy, integrity, and originality of the final manuscript.

Conflict of Interest

The author declares no conflict of interest.

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Statement

The study adhered to all relevant ethical guidelines and was approved by the appropriate ethics committee of Teofilo Kisanji University.

Funding Information

This work received no external funding.

References

1. United Republic of Tanzania. Education sector development programme: secondary education development plan (2004–2009). Dar es Salaam: Ministry of Education and Vocational Training; 2005.

2. Vygotsky LS. Mind in society: the development of higher psychological processes. Cambridge (MA): Harvard University Press; 1978.

3. Bruner JS. The process of education. Cambridge (MA): Harvard University Press; 1960.

4. Rahayuningsih SH, Asrowi, Gunarhadi, Ramli M. Meta-analysis effectiveness of using the discovery learning model on students' 21st-century critical thinking skills. J Penelit Pendidik IPA. 2024;10(10):685-92. doi: 10.29303/jppipa.v10i10.8961

5. Dewey J. Experience and education. New York: Macmillan; 1938.

6. Sharkey S, Weimer M. Learner-Centered Teaching: Five Key Changes to Practice. Teaching Sociology. 2003;31(2):251. doi: https://doi.org/10.2307/3211318

7. Schweisfurth M. Learner-centred education in developing country contexts: From solution to problem?. International Journal of Educational Development. 2011;31(5):425-432. doi: https://doi.org/10.1016/j.ijedudev.2011.03.005

8. Piaget J. Science of education and the psychology of the child. New York: Orion Press; 1970.

9. Tokas R. Freire, Paulo 'Pedagogy of the Oppressed', Continuum Publication, New York, 2005. Lear. Commun. 2015;6(2and3):283. doi:

https://doi.org/10.5958/2231-458x.2015.00027.5

10. Erskine CA, Tomkin A. Evaluation of the effect of the group discussion method in a complex teaching program. Acad Med. 1963;38(12):1036-42.

11. Garside C. Look who's talking: A comparison of lecture and group discussion teaching strategies in developing critical thinking skills. Communication Education. 1996;45(3):212-227. doi: https://doi.org/10.1080/03634529609379050

12. Slavin RE. Cooperative learning: theory, research, and practice. Englewood Cliffs (NJ): Prentice Hall; 1990.

13. Theobald EJ, Hill MJ, Tran E, Agrawal S, Arroyo EN, Behling S, *et al*. Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. Proc Natl Acad Sci U S A. 2020;117(12):6476-83. doi: 10.1073/pnas.1916903117

14. Msonde SE, Msonde CE. Re-innovation of learner-centered pedagogy in Tanzania's secondary schools. J Educ. 2019;199(3):142-54. doi: 10.1177/0022057419854343

15. John G. Implementation of innovative teaching methods on enhancing quality teaching and learning in secondary schools in morogoro tanzania. Pjp. 2024;11(2):264-277. doi: https://doi.org/10.33650/pjp.v11i2.9429

16. Ndomondo E. Instructional resources for innovative history teaching and learning in Tanzania secondary schools: exploring availability and utilization. Cogent Arts & Humanities. 2024;11(1):2382526. doi: https://doi.org/10.1080/23311983.2024.2382526

17. Bharath P. Unravelling second-order concepts in South African history textbooks. Yesterday Today. 2024;(31):10-36. doi: 10.17159/2223-0386/2024/n31a2

18. Bušljeta R. Effective Use of Teaching and Learning Resources. Cphpj. 2013;5(2):55-69. doi: https://doi.org/10.2478/cphpj-2013-0014

19. Simesso MD, Gutu TS, Tarekegn WM. The Contribution of Using Cooperative Learning Methods on Students' Achievement and Retention in Secondary Schools during Chemistry Lesson. Education Research International. 2024;2024:1-12. doi: https://doi.org/10.1155/2024/1830124

20. Kebede YA, Zema FK, Geletu GM, Zinabu SA. Cooperative Learning Instructional Approach and Student's Biology Achievement: A Quasi-Experimental Evaluation of Jigsaw Cooperative Learning Model in Secondary Schools in Gedeo Zone, South Ethiopia. Sage Open. 2025;15(1). doi: https://doi.org/10.1177/21582440251318883

21. Tadesse T, Ware H, Asmare A, Gillies RM. Enhancing Student Engagement and Outcomes: The Effects of Cooperative Learning in an Ethiopian University's Classrooms. Education Sciences. 2024;14(9):975. doi: https://doi.org/10.3390/educsci14090975

22. Ishtiaq M. Book Review Creswell, J. W. (2014). Research Design: Qualitative, Quantitative and Mixed Methods Approaches (4th ed.). Thousand Oaks, CA: Sage. Elt. 2019;12(5):40. doi: https://doi.org/10.5539/elt.v12n5p40

23. Kothari CR. Research methodology: methods and techniques. 2nd ed. New Delhi: New Age International Publishers; 2004.

24. Tanzania Institute of Education. History syllabus for ordinary secondary education, Form I-IV. Dar es Salaam: Tanzania Institute of Education; 2023.

25. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101. doi: https://doi.org/10.1191/1478088706qp063oa

26. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.

Additional Information

How to Cite

APA 7th Edition: Sheyo, H. J. (2026). Group Discussion and Students' Academic Performance in History Among Secondary School Students in Mbeya City, Tanzania: A Descriptive Comparative Study. Education and Learning, 2(2), 67-79. https://doi.org/10.58920/edu0202779

Vancouver: Sheyo HJ. Group Discussion and Students' Academic Performance in History Among Secondary School Students in Mbeya City, Tanzania: A Descriptive Comparative Study. Education and Learning. 2026;2(2):67-79. https://doi.org/10.58920/edu0202779

Harvard: Sheyo, H. J. (2026) 'Group Discussion and Students' Academic Performance in History Among Secondary School Students in Mbeya City, Tanzania: A Descriptive Comparative Study', Education and Learning, 2(2), pp. 67-79. doi: 10.58920/edu0202779

Publisher Note

All claims expressed in this article are solely those of the authors and do not necessarily reflect the views of the publisher, the editors, or the reviewers. Any product that may be evaluated in this article, or claim made by its manufacturer, is not guaranteed or endorsed by the publisher. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License. You may share and adapt the material with proper credit to the original author(s) and source, include a link to the license, and indicate if changes were made.