Performance-based assessment of pre-service teachers' planning skills in selected topics in the chemistry curriculum for colleges of education in Ghana

Anna Mwinbuabu Naah[1]
Valentina Osei-Himah[2]

[1]annanaah2002@yahoomail.com
[2]oseihimah@gmail.com

[1]Offinso College of Education, [2]Atebubu College of Education, [1,2]Ghana

https://doi.org/10.51867/ajernet.7.1.46

ABSTRACT

This study examined the extent to which pre-service teachers in a Ghanaian college of education exhibit planning skills as a key science process skill, using performance-based assessment. Grounded in constructivist learning theory, inquiry-based science education, and authentic assessment, the study adopted a quantitative approach complemented with an alternative assessment strategy. The target population was level 300 pre-service teachers. A purposive sample of 227 level 300 pre-service teachers offering primary education participated in the study. Data were collected using three performance-based tasks designed from the general chemistry curriculum—distillation, density, and osmosis—implemented through a station-by-station assessment model. Descriptive statistics were used to determine the extent of planning skills exhibited, while independent samples t-tests examined gender differences. Results revealed that pre-service teachers demonstrated generally high levels of planning skills across all three tasks, particularly in identifying procedures, sequencing steps, and organising experimental activities. However, persistent weaknesses were observed in the integration of safety measures, measurement accuracy, and explicit conceptual justification. Inferential analysis showed no statistically significant gender differences in planning skills across the tasks. These findings suggest that planning skills develop equitably among male and female pre-service teachers when assessed through authentic, inquiry-oriented tasks. It is recommended that colleges of education explicitly integrate laboratory safety reasoning into the teaching and assessment of science process skills.

Keywords: Ghana, Inquiry-Based Learning, Performance-Based Assessment, Planning Skills, Pre-Service Teachers, Science Process Skills

I. INTRODUCTION

One of the main goals of science education is to give every citizen a fundamental culture structured around subjects of social and personal significance, which will allow students to make comparatively well-founded decisions about scientific and technology matters and to apply their scientific knowledge to everyday life (García-Carmona et al., 2017). For learners to develop scientific skills, it is important for them to be trained in the processes of seeking solutions to problems through scientific investigations and experimentations (National Council for Curriculum and Assessment [NACCA], 2019). Scientific investigations and experimentations allow learners to gain personal experiences of science through hands-on activities and to develop the skills associated with the practice of science (Photo, 2025). These scientific investigations and experimentation are developed through the acquisition of science process skills (Katayev, 2023; Hammann et al., 2008). Science process skills are used to describe a set of broadly transferable abilities that are reflective of what scientists do. They are fundamental to science, allowing everyone to conduct investigation, analyze data gathered, interpret them and draw conclusions. Science process skills tend to last longer than learned content, and it is believed these thinking patterns can be readily transferred to new situations (Dolapcioglu & Subasi, 2022).

Assessment determines what, when and how learners learn; additionally, assessment should enable a teacher have enough information about what learners know, understand and can do and to help them learn (Black & Wiliam, 2018). Assessment in education may be described as the process of gathering, interpreting, recording, using and communicating information about learners' achievements with respect to knowledge, concepts, skills and attitudes (Levy-Feldman, 2025). Assessment is an important part of instructional programmes in schools (Atasoy & Kaya, 2022); as such assessment in science should be broaden to include performance-based assessment that will provide a documentary record or a clear picture of learners' abilities, interests and experiences or skills (Yakob et al, 2021).

One important hurdle science educators must overcome is getting learners to have interest in science, so as to facilitate intellectual development of learners, prepare and empower them to become more actively engaged in the decisions made in science. It is thought that learners' attitudes toward science may have an effect on learners' motivation,

interest, and achievement in the sciences, gender notwithstanding. However, there are various biological differences in the make-up of humans as male and female which in most cases have led to the notion that one sex may have a 'learning edge' over the other sex (Rozgonjuk et al., 2024). Some people believe that because men are regarded as the dominant and even superior sex, they intrinsically have better brains and can learn much better than women (Beroíza-Valenzuela, 2025). It has also been generally acclaimed that girls have better verbal ability than boys, whilst boys have greater visual spatial ability than girls. Perceived differences in ability of learners has often led to branding as noted by Kashikar et al. (2026), thus, it is very easy to label learners as non-performers in science because they tend to perform badly on paper-and-pencil laboratory assessment examinations without taking due cognisance of the possible factors that work against them.

A research work on the use of Science Performance Assessment instrument by Antwi et al. (2021), indicated physics teachers need to expose their learners more to non-traditional laboratory activities in order to develop and improve upon their skills of planning, performing and reasoning. Ogunsola-Bandele and Kennepohl (2024) attest that, alternative assessment approaches which include non-traditional, performance-based laboratory tasks provide a clearer picture of learners' actual laboratory competencies and understanding than traditional assessments alone.

In Ghana, several researchers have conducted performance-based assessments and drawn interesting conclusions. Koomson (2020) discovered that learners' abilities to infer, predict, classify, and use other integrated science process skills were lacking. Koomson et al. (2024) concluded that "learners could demonstrate some degree of interpreting, inferring and predicting skills. Antwi et al. (2021) concluded that non-traditional tasks are better able to measure learners' laboratory skills. Johnson (2017) observed that learners could demonstrate some level of proficiency in the skills of planning and performing. However, performance-based assessment of science planning skills in relation to pre-service teachers at Ghanaian colleges of education is understudied, despite the variety of research on science instruction in Ghana using science performance assessment instruments to evaluate learners' process skills. This underscores the need for this study.

1.1 Statement of the Problem

In Ghana, the University of Coast General Chemistry Curriculum for colleges of education of 2019 emphasizes not only knowledge acquisition and comprehension but also the development of experimental and science process skills to help students understand the natural world and approach problems logically. Despite these objectives, many students struggle to demonstrate clear understanding of practical knowledge and scientific concepts, suggesting gaps between curriculum intentions and actual student performance in science learning. Additionally, research indicates that active engagement in learning, modelled after how scientists work, enhances students' understanding of scientific concepts, science process skills, and the nature of science itself, benefiting both academic success and personal development (Ogunsola-Bandele & Kennepohl, 2024). However, studies on gender differences in science performance show mixed results, with some reporting differences between males and females (Rozgonjuk et al., 2024; Beroíza-Valenzuela, 2025), while others found none (Antwi et al., 2021; Johnson, 2017).

1.2 Research Questions

The study was aimed at finding answers to the following research questions:

i. To what extent do pre-service teachers of a college of education offering general chemistry exhibit planning skills?
ii. Which gender shows more proficiency in planning skills?

II. LITERATURE REVIEW

2.1 Theoretical Review

This study is anchored in the Constructivist Inquiry–Planning Skills Development Theory, grounded in the seminal works of Jean Piaget and Jerome Bruner. The theory provides a lens for examining how pre-service teachers develop and exhibit planning skills in general chemistry through active knowledge construction, discovery-oriented inquiry, and authentic performance-based assessment.

2.1.1 Constructivist Inquiry–Planning Skills Development Theory

Constructivist learning theory posits that learners actively construct knowledge through interaction with their environment rather than passively receiving information. Learning is most effective when learners engage in meaningful tasks that require them to apply prior knowledge, test ideas, and reflect on outcomes (Jumaah, 2024). In science education, this perspective aligns strongly with inquiry-based learning, where learners investigate problems, design experiments, and construct scientific understanding through hands-on experiences. Within this framework, planning skills are conceptualised as a core science process skill, encompassing the ability to analyze scientific problems, formulate hypotheses, identify and control variables, select appropriate experimental procedures, choose suitable

apparatus, and organise data collection strategies. These skills are foundational to scientific inquiry and essential for the development of scientific literacy.

The framework posits that engagement in inquiry-based science tasks, when evaluated through performance-based assessment, leads to the development and accurate measurement of pre-service teachers' planning skills. Performance-based assessment provides authentic contexts that allow learners to demonstrate procedural and cognitive skills that are often inadequately captured by traditional paper-and-pencil tests (Vlachopoulos & Makri, 2024). By observing learners as they plan and design experiments, assessors gain richer evidence of learners' reasoning, decision-making, and understanding of scientific processes.

The framework is explicitly aligned with Ghana's Standards-Based Curriculum for Science (NACCA, 2019), which emphasises inquiry-based learning, competency development, and authentic assessment. The curriculum identifies planning, controlling variables, experimentation, and data interpretation as essential competencies for scientific literacy and lifelong learning. Performance-based assessment within this framework reflects the curriculum's call for learner-centred assessment approaches that integrate teaching, learning, and assessment. Gender is incorporated as a moderating variable in the framework to examine potential differences in the exhibition of planning skills among male and female pre-service teachers. The framework therefore explains how inquiry-oriented instruction and authentic assessment interact to support the development, exhibition, and measurement of planning skills necessary for effective science teaching at the basic school level.

2.2 Empirical Review

Empirical studies consistently highlight the importance of planning skills in science education and their role in promoting scientific literacy. Planning skills are central to conducting scientific investigations, as they involve hypothesis formulation, method selection, variable control, and outcome interpretation (Temiz, 2020). These skills enable learners to apply theoretical knowledge to real-world contexts, fostering deeper conceptual understanding. Research indicates that assessing planning and experimental skills poses significant challenges. Hyytinen et al. (2021) argue that planning effective experiments requires mastery of multivariable reasoning - particularly control of variables - which is difficult to measure using traditional multiple-choice assessments. Their findings demonstrate that performance tasks and process-based assessments provide richer evidence of learners' strategic planning and causal reasoning than selected-response formats.

Goles-Sabellina (2021) emphasises that performance-based assessment supports the assessment of a wide range of learning outcomes, including conceptual understanding, psychomotor skills, communication, problem-solving, and teamwork. Observational assessment of learners' work enhances reflection, self-assessment, and expression of thinking processes (Yaghoubi et al., 2025). Wren (2019) further outlines that effective performance-based assessment involves defining assessment purpose, selecting appropriate tasks, and developing clear scoring criteria through rubrics. Learners' proficiency in planning skills remains comparatively low among science process skills, particularly in tasks involving experimental design, control of variables, and multivariable reasoning. Analyzes of inquiry performance data show that learners often struggle to independently formulate coherent investigation plans and align procedures with research questions (Teig, 2024). These difficulties are attributed partly to the abstract and integrative nature of planning skills, which require higher-order cognitive coordination.

Further research demonstrates that traditional selected-response assessments inadequately capture planning competence, often underestimating learners' procedural and strategic reasoning (Hyytinen et al., 2021). In contrast, performance-based and inquiry-oriented assessments provide richer evidence of students' planning abilities by allowing them to design, justify, and revise experimental procedures. Inquiry-based and project-based instructional approaches significantly improve students' planning skills when adequate scaffolding and feedback are provided (García-Carmona et al., 2017). Overall, current evidence highlights a persistent gap between curriculum expectations and learners' demonstrated planning proficiency, underscoring the need for authentic assessment practices. Studies focusing on inquiry-based science education reveal strong links between inquiry-oriented teaching and the development of planning skills. García-Carmona et al. (2017) found that learners with strong planning abilities are better equipped to design experiments and connect scientific theory with practice. Their research supports the view that inquiry-based instruction fosters the gradual acquisition of experimental design skills through structured and reflective planning activities. Temiz (2020), using scenario-based assessments, concluded that science process skills are both intellectual and physical skills essential for inquiry-based science learning. However, despite curriculum emphasis on these skills, evidence suggests that many learners struggle to demonstrate practical and procedural understanding of scientific concepts.

Recent empirical research on gender differences in planning and science process skills presents mixed findings. Some studies suggest that male students outperform female students in aspects of scientific inquiry, particularly in experimental design and overall science process skills at the junior secondary level, indicating a possible gender-related advantage in structured planning tasks (Nicol et al., 2022; Teig, 2024). Conversely, other research reports no significant gender differences in planning and designing experiments following guided inquiry-based instruction, highlighting that

instructional context and scaffolding can mitigate gender disparities (Schoenfeld, 2023). Additional studies show that while males may report greater involvement in procedural planning tasks, females often excel in cognitive planning aspects and organisation (Alalouch, 2021). Overall, the literature suggests that gender effects on planning skills are context-dependent, influenced by instructional methods, assessment approaches, and the specific nature of the planning tasks in scientific inquiry.

Within the context of pre-service teacher education, educators play a critical role in shaping learners' acquisition of science process skills. Pre-service teachers in Ghana are trained to teach science at the basic school level, and the quality of their training directly influences curriculum implementation. The dependent variable in this study is planning skills, defined as a core science process skill demonstrated by pre-service teachers during performance-based chemistry tasks. Planning skills involve analysing scientific problems, designing experimental procedures, organising materials, sequencing steps, controlling variables, and integrating safety and measurement considerations. These skills were operationalised through authentic tasks in distillation, density determination, and osmosis, consistent with inquiry-based and constructivist perspectives on science learning. Planning skills align with Ghana's Standards-Based Science Curriculum as key indicators of scientific literacy and instructional readiness. The independent variable was the performance-based assessment approach, implemented through station-based laboratory tasks, which research shows elicit planning and experimental design skills more effectively than traditional assessments Research indicates that pre-service teachers who develop strong cognitive, affective, and psychomotor skills are better positioned to foster inquiry-oriented learning and scientific literacy among their future learners. Despite the acknowledged importance of science process skills, there remains limited empirical evidence on how pre-service teachers exhibit planning skills during their formative training years, particularly using performance-based assessment. This gap justifies the present study, which assesses the extent to which pre-service teachers demonstrate planning skills through authentic chemistry-based performance tasks and examines gender differences in skill exhibition. The study therefore contributes to science education research by providing evidence to inform teacher education practices, assessment strategies, and competency-based curriculum implementation in Ghana.

III. METHODOLOGY

3.1 Research Design

The study employed a quantitative research design, complemented by an alternative assessment approach, to provide a comprehensive evaluation of pre-service teachers' planning skills in science. Quantitative methods allowed for systematic measurement and statistical comparison of performance, while the alternative assessment approach, specifically performance-based assessment, provided authentic and learner-centered evaluation of participants' ability to engage in scientific inquiry. Performance-based assessments allow teachers to evaluate how learners organize, plan, execute, analyze, and communicate scientific investigations, simulating real-world scientific practices (Darling-Hammond et al., 2020). The study used a station–by–station performance model, where tasks were designed to progressively assess different aspects of planning skills. Participants rotated through three distinct stations, completing one task per station, ensuring that each student had the opportunity to demonstrate their competencies in planning and performing experiments. The station model enhances reliability by minimizing the influence of proximity and peer observation but may not fully capture the continuous manipulation skills involved in laboratory experiments, as the activities are artificially segmented (Organisation for Economic Co-operation and Development [OECD], 2023).

3.2 Study Area

The study was conducted among pre-service teachers in a college of education in Ghana, in level 300 and preparing for Supported Teaching in Schools (STS), commonly referred to as macro teaching practice. The pre-service teachers are being trained to teach both science (Basic 1–3) and integrated science (Basic 4–6) in basic schools. Performance-based tasks were conducted in lecture halls or laboratory spaces equipped with the necessary apparatus to simulate authentic science teaching and learning environments. The study setting allowed for controlled assessment conditions while reflecting realistic classroom and laboratory contexts, ensuring ecological validity.

3.3 Target Population

The population consisted of 417 level 300 pre-service teachers preparing for STS. Of these, 227 were enrolled in primary education programs and were purposively selected for the study. This focus was chosen because these participants were responsible for teaching science at the basic school level, making their performance in planning and performing scientific tasks directly relevant to classroom teaching outcomes. The pre-service teachers were preparing for Supported Teaching in Schools (STS), commonly referred to as macro teaching practice. The pre-service teachers were trained to teach both science (Basic 1–3) and integrated science (Basic 4–6) in basic schools.

3.4 Sampling Procedures and Sample Size

A purposive sampling technique was used to select 227 primary education pre-service teachers from the larger population. Purposive sampling was appropriate because the study required participants with prior knowledge of the science topics under investigation, ensuring that all participants could engage meaningfully with the performance-based tasks. The selection criteria included participants' enrollment in primary education programs and their readiness to teach science content at the basic school level.

3.5 Data Collection Instruments and Procedures

Data collection was conducted using three performance-based tasks, each carefully designed to assess planning skills in general chemistry. The tasks were grounded in concepts that have practical relevance and real-life application, such as: *Transport systems in organisms, measuring densities of irregular objects* and *Water purification processes.* The station–by–station model was used to organize the assessment. Tasks were placed on desks in a sequential order, and participants moved from one station to another, completing a unique, independent task at each station. This ensured that the performance on one task did not influence subsequent tasks. Pre-service teachers were required to:

- Analyze a scientific problem
- Plan a relevant experiment
- Execute the experiment
- Collect and organize data
- Analyze results
- Communicate findings effectively

This approach mirrors authentic scientific practices, enabling the assessment of both cognitive and practical skills in planning and execution.

A pre-test was conducted to ensure the validity and reliability of the instruments. Participants were asked to provide feedback on the difficulty of tasks, familiarity with apparatus, clarity of instructions, and relevance of concepts. Based on their feedback, the tasks were revised to improve clarity and appropriateness for the actual study. An opinionnaire was used during pre-testing but not included in the main study. Each task had a scoring rubric to ensure systematic and objective evaluation. The rubrics allowed for assessment of multiple dimensions of performance, including accuracy, organization, logical reasoning, and procedural competence.

3.6 Data Analysis

Data analysis was conducted using both descriptive and inferential statistics. Descriptive statistics were employed to answer the first research question: *To what extent do pre-service teachers exhibit planning skills?* Measures such as means, standard deviations, and percentages were used to summarize the level of proficiency demonstrated by participants across the three tasks. Inferential statistics, specifically independent t-tests, were used to address the second research question: *Are there significant differences between male and female pre-service teachers in planning skills?* The t-tests allowed for comparison of mean scores across gender groups, providing insight into potential gender differences in the acquisition of planning skills.

IV. FINDINGS & DISCUSSION

4.1 Findings

Research Question 1

To what extent do Pre-Service Teachers of a College of Education Offering General Chemistry Exhibit Planning Skills?

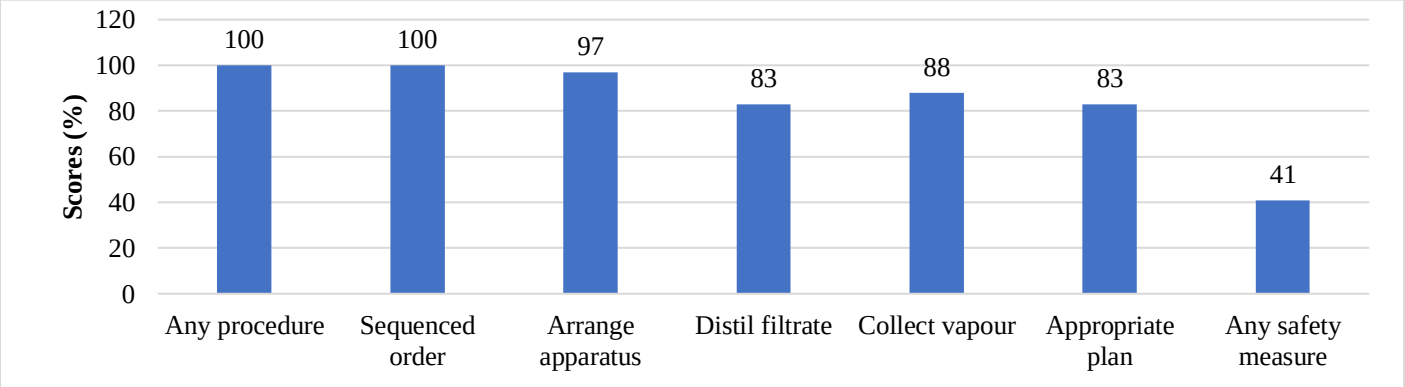

Figure 1

Proficiency Level: Assessment of Planning Skills for Task 1 (Distillation)

In general, pre-service teachers proficiency level was higher than 80% for almost all the skills exhibited for task '1' (distillation) as shown in figure 1. Most of the pre-service teachers exhibited high level of planning skill in conducting the experiment related to distillation. The pre-service teachers could plan the experiment by coming-out with the necessary process skills required to execute the task. However, with all the high level of proficiency exhibited, most of the pre-service teachers could not state any 'safety measure' to guide the execution of the task.

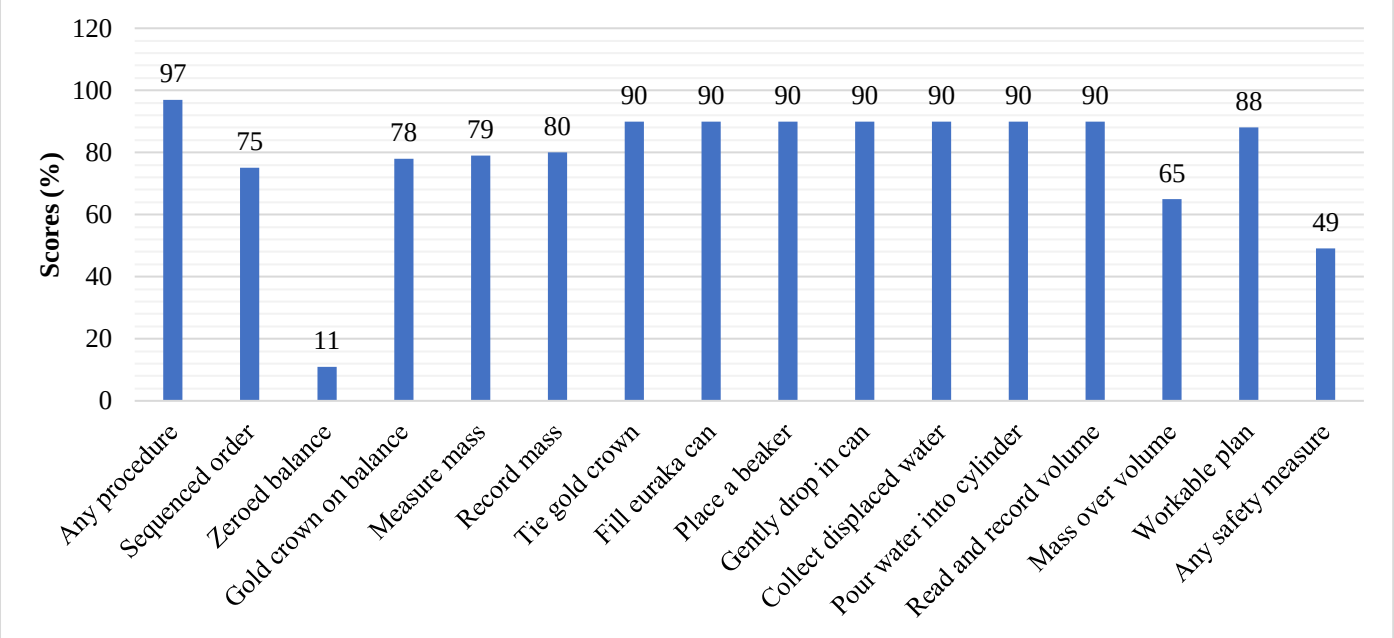

Figure 2
Proficiency Level: Assessment of Planning Skills for Task 2 (Density)

The pre-service teachers generally exhibited proficiency level higher than 65% for almost all the skills exhibited for task '2' (density) as shown in figure 2. Al most of the pre-service teachers exhibited high level of planning skill in conducting the experiment related to density. They could plan the experiment by coming-out with the various process skills required to execute the task. However, with all the high level of proficiency exhibited by pre-service teachers almost all the pre-service teachers could not 'zero balance', and few could not state the formula 'mass over volume' and less than 50% could not state any 'safety measure' required to guide the execution of the task.

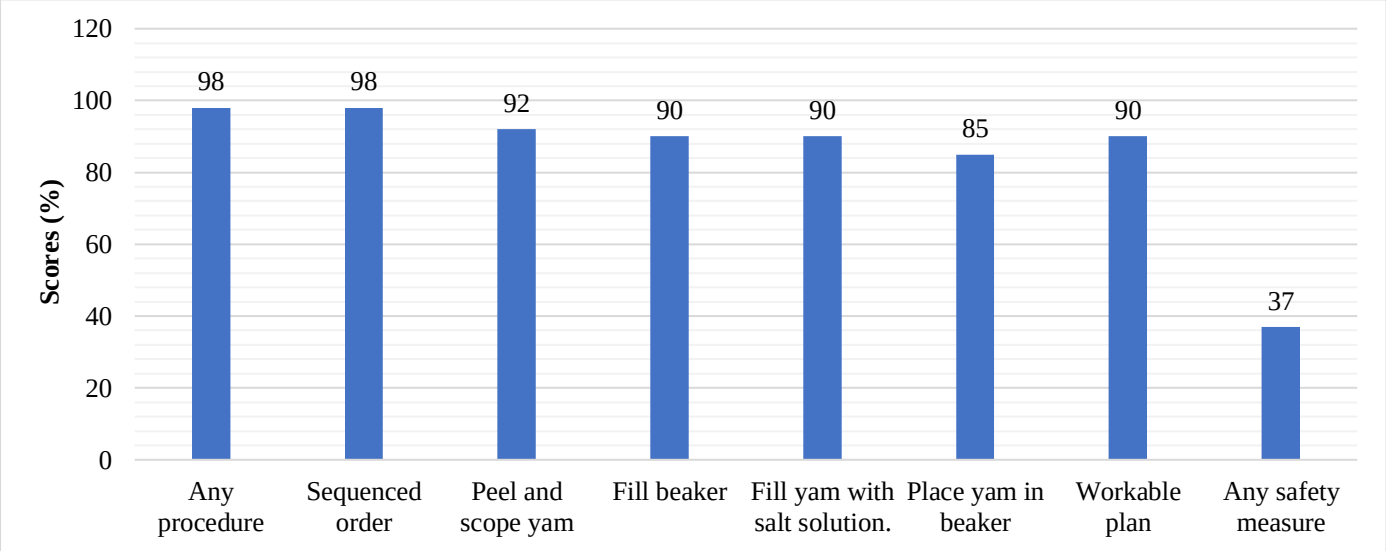

Figure 3
Proficiency Level: Assessment of Planning Skills for Task 3 (Osmosis)

Generally, the pre-service teachers exhibited high proficiency level above 80% for almost all the skills exhibited for task '3' (osmosis) as shown in figure 3. Al most all the pre-service teachers exhibited high level of planning skill in

conducting the experiment related to osmosis. The pre-service teachers could plan the experiment by coming-out with the various process skills required to execute the task. However, with all the high level of proficiency exhibited by pre-service teachers almost all the pre-service teachers could not state any 'safety measure' required to guide the execution of the task.

The results indicate that pre-service teachers demonstrated generally high levels of planning skills across the three performance-based tasks - distillation, density, and osmosis - suggesting a strong capacity to organise experimental procedures, identify variables, and sequence investigative steps logically. This overall finding is consistent with the results of Katayev (2023) and Dolapcioglu and Subasi (2022) who attest that engagement in practical science activities enhances learners' procedural and inquiry-related competencies. From a constructivist perspective, such competencies are developed when learners actively construct knowledge through interaction with authentic tasks and materials rather than through passive instruction (García-Carmona et al., 2017).

Task 1: Distillation: For the distillation task, proficiency levels exceeded 80% for almost all planning-related skills, indicating that most pre-service teachers were able to design appropriate experimental procedures, select relevant apparatus, and identify the sequence of steps required to carry out the task effectively. Distillation is a core topic in general chemistry and closely linked to everyday experiences, which may have facilitated the transfer of theoretical knowledge into practical planning. This finding aligns with earlier studies suggesting that familiarity with content and context supports the development of science process skills, particularly planning and procedural reasoning (Temiz, 2020; García-Carmona et al., 2017).

However, despite the high overall proficiency, most pre-service teachers failed to articulate safety measures relevant to the experiment. This omission suggests that their conception of planning skills was largely procedural, with limited attention to risk assessment and laboratory safety. Empirical studies have similarly reported that learners often overlook safety considerations during experimental planning unless these are explicitly emphasised in instruction and assessment criteria (Hammann et al., 2008; Katayev, 2023). This finding highlights a critical gap in the holistic development of planning skills, as safety reasoning is a fundamental component of scientific practice.

Task 2: Density: In the density task, pre-service teachers again demonstrated relatively high proficiency, with most skills exceeding 65%. Participants were generally able to plan the experimental steps needed to determine density, including identifying materials and outlining procedures. This supports previous research indicating that performance-based tasks are effective in eliciting learners' ability to apply scientific knowledge in practical contexts (Goles-Sabellina, 2021).

Nonetheless, important weaknesses were evident. Almost all participants failed to indicate the need to zero the balance, some could not correctly state the density formula (mass ÷ volume), and fewer than half identified relevant safety measures. These results suggest that while procedural sequencing was present, attention to measurement accuracy and conceptual precision was inconsistent. Similar findings have been reported in studies showing that learners may demonstrate surface-level planning skills without fully integrating conceptual understanding and metacognitive control (Dolapcioglu & Subasi, 2022; García-Carmona et al., 2017). This reinforces the view that effective planning requires the coordinated use of conceptual knowledge, procedural knowledge, and reflective judgment.

Task 3: Osmosis: Results from the osmosis task revealed proficiency levels above 80% for nearly all planning skills, indicating strong performance in organising experiments related to biological transport processes. Osmosis is commonly taught through demonstrations and simple investigations at the basic school level, which may explain participants' confidence and competence in planning this task. Inquiry-based exposure to such concepts has been shown to promote learners' ability to design experiments and anticipate outcomes (Temiz, 2020; Photo, 2025).

Despite this strong performance, the persistent failure to specify safety measures mirrors the pattern observed in the distillation and density tasks. Even though osmosis experiments are often perceived as low-risk, the absence of safety considerations suggests that pre-service teachers may not conceptualise safety as an integral element of experimental planning. This supports earlier assertions that learners tend to treat safety as implicit unless assessment tasks explicitly require its consideration (Hammann et al., 2008; Katayev, 2023).

Cross-Task Interpretation: Across all three tasks, a consistent pattern emerged: pre-service teachers demonstrated strong procedural planning skills but weak integration of safety awareness, measurement accuracy, and explicit conceptual justification. This suggests that planning skills were developed primarily as step-by-step procedural competencies rather than as comprehensive scientific practices that include safety, precision, and critical evaluation. From a constructivist standpoint, this may reflect learning experiences that emphasise task completion over reflective engagement with the underlying principles guiding experimental decisions (García-Carmona et al., 2017).

Importantly, these findings also illustrate the value of performance-based assessment in revealing aspects of learners' planning skills that are unlikely to be captured through traditional paper-and-pencil tests. By requiring pre-service teachers to plan experiments "from scratch," the assessment exposed omissions and misconceptions related to safety and measurement that might otherwise remain hidden (Hammann et al., 2008; Goles-Sabellina, 2021).

Overall, while the high proficiency levels observed across tasks indicate promising preparation among pre-service teachers, the recurring gaps underscore the need for teacher education programmes to place greater emphasis on safety reasoning, measurement validity, and explicit articulation of planning decisions. Embedding these elements within inquiry-based instruction and performance-based assessment practices may support the development of more robust planning skills and better prepare pre-service teachers for effective implementation of competency-based science education at the basic school level.

Research Question 2

Which Gender shows more Proficiency in Planning Skills?

Table 1

Distribution of Mean, Standard Deviation and Standard Error Mean by Gender on Task 1

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Any procedure	male	137	1.0000	.00000[a]	.00000
	female	98	1.0000	.00000[a]	.00000
Sequence order	male	137	1.0000	.00000[a]	.00000
	female	98	1.0000	.00000[a]	.00000
Arranged apparatus	male	137	.9635	.18821	.01608
	female	98	.9694	.17315	.01749
Distilled filtrate	male	137	.8467	.36158	.03089
	female	98	.8061	.39737	.04014
Collected vapour	male	137	.8832	.32235	.02754
	female	98	.8776	.32949	.03328
Appropriate plan	male	137	.8248	.38152	.03260
	female	98	.8367	.37151	.03753
Any safety measure	male	137	.4307	.49699	.04246
	female	98	.3878	.48974	.04947

Table 2

Distribution of p and t Values by Gender of Pre-service Teachers on Task 1

		Levene's Test for Equality of Variances			
		F	Sig.	t	df
Arranged apparatus	Equal variances assumed				
	Equal variances not assumed			-.248	218.786
Distilled filtrate	Equal variances assumed	2.617	.107	.814	233
	Equal variances not assumed			.801	196.714
Collected vapour	Equal variances assumed	.069	.793	.132	233
	Equal variances not assumed			.131	206.314
Appropriate plan	Equal variances assumed	.229	.632	-.239	233
	Equal variances not assumed			-.240	212.358
Any safety measure	Equal variances assumed	1.777	.184	.656	233
	Equal variances not assumed			.658	210.914

The results for Task 1 provide insight into how male and female pre-service teachers exhibited planning skills related to science process skills during the distilation experiment, as well as whether gender differences were statistically meaningful. From table 2, Levene's Test for Equality of Variances was not significant for all variables ($p > .05$), indicating that the assumption of equal variances was met. Therefore, the results from the "equal variances assumed" rows were used for interpretation. The independent samples t-test results show that none of the observed mean differences between males and females were statistically significant. Since all t-values are small and non-significant, the results indicate that gender did not have a significant effect on students' performance in the assessed practical skills.

The finding that there were no statistically significant gender differences in pre-service teachers' exhibition of planning-related science process skills aligns with a growing body of literature suggesting that when learners are engaged in authentic, inquiry-based, and performance-oriented tasks, gender disparities in science achievement tend to diminish (García-Carmona et al., 2017; Katayev, 2023). This result supports the view that science process skills - such as planning experiments, organising procedures, and sequencing investigative steps - are learnable competencies that develop through structured practice rather than innate ability. Consistent with Temiz (2020) and Dolapcioglu and Subasi

(2022), the present study demonstrates that both male and female pre-service teachers were equally capable of applying scientific reasoning to practical contexts when assessed through performance-based methods. From the perspective of constructivist and inquiry-based science education, this suggests that meaningful engagement with hands-on investigations allows all learners to construct scientific understanding through experience, reflection, and problem-solving, irrespective of gender. Importantly, these findings align strongly with Ghana's Standards-Based Curriculum for Science, which emphasises inquiry, experimentation, and the development of transferable science process skills as foundations for scientific literacy (NACCA, 2019). By demonstrating comparable planning skills across genders, the results reinforce the curriculum's commitment to equity, inclusion, and competency-based learning, ensuring that every learner acquires the skills needed to apply science to real-life situations.

In addition, the findings support the goals of Education for Sustainable Development (ESD), which seek to empower learners with critical thinking, problem-solving, and decision-making skills necessary for addressing societal and environmental challenges (García-Carmona et al., 2017). The use of performance-based assessment in this study reflects ESD principles by valuing authentic learning, collaboration, and practical competence over rote memorisation. Thus, the absence of gender differences in planning skills suggests that well-designed, inquiry-driven science instruction and assessment - consistent with Ghana's curriculum and ESD goals - can promote inclusive scientific literacy, preparing pre-service teachers to foster equitable and sustainable science learning in basic schools.

Table 3

Distribution of Mean, Standard Deviation and Standard Error Mean by Gender on Task 2

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Any procedure	male	137	.9489	.22100	.01888
	female	98	.9898	.10102	.01020
Sequence order	male	137	.9343	.24865	.02124
	female	98	.9796	.14212	.01436
Zeroed balance	male	137	.1022	.30401	.02597
	female	98	.1327	.34094	.03444
Gold crown on balance	male	137	.7810	.41507	.03546
	female	98	.7857	.41244	.04166
Measure mass	male	137	.7883	.41000	.03503
	female	98	.7857	.41244	.04166
Record mass	male	137	.8029	.39925	.03411
	female	98	.7959	.40510	.04092
Tie gold crown	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Fill Eureka can	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Place a beaker	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Gently drop in can	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Collect displaced water	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Pour water into cylinder	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Read and record the volume	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Mass over plan	male	137	.6642	.47399	.04050
	female	98	.6327	.48456	.04895
Workable plan	male	137	.8686	.33906	.02897
	female	98	.8980	.30426	.03073
Any safety measure	male	137	.4891	.50171	.04286
	female	98	.4796	.50215	.05073

Table 4

Distribution of p and t Values by Gender of Pre-service Teachers on Task 2

		Levene's Test for Equality of Variances			
		F	**Sig.**	**t**	**df**
Any procedure	Equal variances assumed				
	Equal variances not assumed			-1.905	202.790
Sequence order	Equal variances assumed	11.195	.001	-1.623	233
	Equal variances not assumed			-1.766	223.288
Zeroed balance	Equal variances assumed	2.060	.153	-.720	233
	Equal variances not assumed			-.706	193.963
Gold crown on balance	Equal variances assumed	.029	.864	-.086	233
	Equal variances not assumed			-.086	209.890
Measure mass	Equal variances assumed	.009	.924	.048	233
	Equal variances not assumed			.048	208.351
Record mass	Equal variances assumed	.069	.793	.132	233
	Equal variances not assumed			.131	207.259
Tie gold crown	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Fill Eureka can	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Place a beaker	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Gently drop in can	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Collect displaced water	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Pour water into cylinder	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Read and record the volume	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Mass over plan	Equal variances assumed	.955	.329	.499	233
	Equal variances not assumed			.497	206.296
Workable plan	Equal variances assumed	1.899	.169	-.682	233
	Equal variances not assumed			-.695	221.317
Any safety measure	Equal variances assumed	.082	.775	.142	233
	Equal variances not assumed			.142	208.987

The results for Task 2 provide insight into how male and female pre-service teachers exhibited planning-related science process skills during the density experiment, as well as whether gender differences were statistically meaningful. From table 4, Levene's test indicated unequal variances for some variables (e.g., *Any procedure*, *Sequence order*, and several volume-measurement steps), but the appropriate adjusted t-values were considered in interpretation. Across all assessed skills, none of the t-tests revealed statistically significant gender differences (all |t| values were small and $p > .05$). Even where females recorded slightly higher mean scores (e.g., sequence order, tying the gold crown, and volume-related steps), these differences were not statistically significant. Similarly, males' marginally higher performance on *Mass over plan* and *Any safety measure* did not reach significance.

The findings indicate that gender did not significantly influence pre-service teachers' planning skills in the density task. Both male and female participants demonstrated strong procedural planning abilities, particularly in sequencing steps and handling apparatus, but shared similar weaknesses in areas such as zeroing the balance, explicitly applying the density formula, and integrating safety measures into their plans.

The Task 2 findings indicate that pre-service teachers of both genders demonstrated comparable and generally high levels of procedural planning skills in the density experiment, with no statistically significant gender differences across the assessed indicators. This outcome supports existing literature which suggests that when learners engage in authentic, hands-on, and inquiry-oriented tasks, science process skills such as planning, sequencing procedures, and organising investigations tend to develop in a gender-neutral manner (García-Carmona et al., 2017; Katayev, 2023). Consistent with Temiz (2020) and Dolapcioglu and Subasi (2022), the results show that pre-service teachers were largely

able to apply procedural knowledge - such as measuring mass, collecting displaced water, and recording observations - within a real-world scientific context.

However, the uniformly low performance on items such as *zeroing the balance*, partial difficulty with explicitly stating the density formula (mass ÷ volume), and the moderate attention to safety measures suggest that planning was often conceptualised as a sequence of actions rather than a reflective scientific practice that integrates accuracy, conceptual understanding, and risk awareness. Similar patterns have been reported in prior studies, which note that learners may demonstrate surface-level planning competence while struggling with measurement precision and conceptual justification unless these aspects are explicitly foregrounded in instruction and assessment (Dolapcioglu & Subasi, 2022; Hyytinen et al, 2021).

From the perspective of Ghana's Standards-Based Curriculum for Science, these findings are particularly significant, as the curriculum emphasises inquiry, controlling variables, accurate measurement, and competency-based assessment as core requirements for scientific literacy and lifelong learning (NACCA, 2019). The generally strong procedural performance observed in this study aligns with these curricular expectations, while the identified gaps highlight areas requiring targeted pedagogical attention in teacher education programmes. Moreover, the results resonate with the goals of Education for Sustainable Development (ESD), which advocate for the development of critical thinking, problem-solving, and informed decision-making skills that learners can transfer to everyday life (García-Carmona et al., 2017).

By revealing both strengths and weaknesses in pre-service teachers' planning skills, the performance-based assessment used in Task 2 exemplifies an ESD - aligned approach to assessment - one that promotes reflection, authenticity, and continuous improvement rather than rote performance. Overall, the Task 2 results underscore the need for teacher preparation programmes in Ghana to strengthen the integration of measurement accuracy, conceptual clarity, and safety reasoning within inquiry-based and performance-based science instruction, in order to fully realise the curriculum's vision of equitable, competency-driven, and sustainable science education.

Table 5

Distribution of Mean, Standard Deviation and Standard Error Mean by Gender on Task 3

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Any procedure	male	137	.9781	.14689	.01255
	female	98	.9898	.10102	.01020
Sequence order	male	137	.9781	.14689	.01255
	female	98	.9898	.10102	.01020
Peeled and scoop	male	137	.9124	.28374	.02424
	female	98	.9286	.25886	.02615
Fill beaker with water	male	137	.8613	.34689	.02964
	female	98	.9490	.22117	.02234
Fill yam with salt solution	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Place yam in beaker	male	137	.8102	.39357	.03362
	female	98	.9082	.29028	.02932
Appropriate plan	male	137	.8759	.33089	.02827
	female	98	.9286	.25886	.02615
Any saftey measure	male	137	.3358	.47399	.04050
	female	98	.3980	.49199	.04970

From table 6, Levene's test indicated that the assumption of equal variances was met for some variables but violated for others (e.g., *Fill beaker with water* and *Place yam in beaker*). Consequently, the appropriate rows were considered for interpretation. Despite some statistically significant t-values for specific procedural steps (e.g., *Fill beaker with water* and *Place yam in beaker*), the overall pattern shows that gender differences were not systematic and did not translate into a meaningful overall advantage for either group. The Task 3 results indicate that pre-service teachers demonstrated strong planning skills in the osmosis task regardless of gender, particularly in identifying procedures, sequencing steps, and preparing materials. While females showed slightly higher mean scores on a few procedural items, these differences were limited in scope and did not substantially affect overall planning performance. The consistently low attention to safety measures across both genders highlights a shared weakness in integrating safety considerations into experimental planning.

Table 6

Distribution of p and t Values by Gender of Pre-service Teachers on Task 3

		Levene's Test for Equality of Variances			
		F	**Sig.**	**t**	**df**
Any procedure	Equal variances assumed	1.878	.172	-.681	233
	Equal variances not assumed			-.723	232.679
Sequence order	Equal variances assumed	1.878	.172	-.681	233
	Equal variances not assumed			-.723	232.679
Peeled and scoop	Equal variances assumed	.805	.371	-.446	233
	Equal variances not assumed			-.453	219.656
Fill beaker with water	Equal variances assumed	21.718	.000	-2.201	233
	Equal variances not assumed			-2.362	230.243
Fill yam with salt solution	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Place yam in beaker	Equal variances assumed	19.446	.000	-2.090	233
	Equal variances not assumed			-2.195	232.765
Appropriate plan	Equal variances assumed	7.244	.008	-1.314	233
	Equal variances not assumed			-1.367	231.093
Any saftey measure	Equal variances assumed	3.416	.066	-.976	233
	Equal variances not assumed			-.970	204.323

The findings from Task 3 indicate that pre-service teachers demonstrated high levels of planning skills in the osmosis experiment, with both male and female participants showing strong ability to identify appropriate procedures, sequence experimental steps, and organise materials. This supports the view that engagement in hands-on scientific investigations enhances learners' science process skills, particularly planning, which is foundational to scientific literacy (Katayev, 2023; Dolapcioglu & Subasi, 2022). In line with García-Carmona et al. (2017), the ability of pre-service teachers to translate theoretical knowledge of osmosis into coherent experimental plans reflects the broader goal of science education: enabling learners to apply scientific understanding to real-life contexts and make informed decisions.

The generally high proficiency observed aligns with Ghana's Standards-Based Curriculum for Science, which emphasises inquiry-based learning, experimentation, and the development of transferable science process skills such as planning, controlling variables, and interpreting results (NACCA, 2019). From an Education for Sustainable Development (ESD) perspective, these competencies are critical, as they promote problem-solving, critical thinking, and responsible decision-making needed for addressing real-world challenges. However, the consistently low attention to safety measures observed in Task 3 suggests that planning was largely conceived as procedural sequencing rather than as a holistic scientific practice that integrates risk awareness and ethical responsibility. Similar concerns have been raised in the literature, where learners often omit safety considerations unless they are explicitly embedded in instructional and assessment frameworks (Hammann et al., 2008; Katayev, 2023). This gap highlights the need for teacher education programmes in Ghana to strengthen the explicit integration of safety reasoning within inquiry-based and performance-based assessments, in order to fully realise the curriculum's vision of competency-based, learner-centred, and sustainable science education.

V. CONCLUSION & RECOMMENDATIONS

5.1 Conclusion

This study set out to assess the extent to which pre-service teachers in a Ghanaian College of Education exhibit planning skills as a key science process skill, using performance-based assessment within authentic practical contexts. Drawing on constructivist learning theory, inquiry-based science education, and authentic assessment frameworks, the findings provide important insights into the strengths and gaps in pre-service teachers' preparedness to implement competency-based science instruction at the basic school level. Overall, the results indicate that pre-service teachers demonstrated relatively high levels of planning proficiency across the three performance tasks—distillation, density, and osmosis—suggesting a strong capacity to organise experimental procedures, identify relevant materials, and sequence investigative steps logically.

Across tasks, proficiency levels exceeded 80% for most planning-related skills, particularly in selecting appropriate procedures, arranging apparatus, and identifying the correct sequence of actions. These findings affirm the value of performance-based assessment in eliciting learners' ability to apply theoretical knowledge to practical problem-solving situations, consistent with earlier research emphasising the durability and transferability of science process skills compared to rote content knowledge From a constructivist perspective, the observed proficiency reflects learning

experiences that allowed pre-service teachers to actively construct understanding through engagement with authentic tasks rather than passive recall.

However, despite these strengths, a persistent and critical weakness emerged across all three tasks: the near-universal failure of pre-service teachers to explicitly state appropriate safety measures during experimental planning. This omission was evident even in tasks perceived as low-risk, such as osmosis, indicating that safety considerations were not fully integrated into participants' conceptualisation of planning skills. Similarly, in the density task, gaps were observed in measurement accuracy (e.g., failure to indicate zeroing the balance) and conceptual precision (e.g., omission of the density formula). These findings suggest that while pre-service teachers possess procedural planning abilities, their planning skills remain largely procedural rather than holistic, with limited integration of safety reasoning, measurement validity, and reflective justification. This pattern aligns with prior studies reporting that learners often demonstrate surface-level competence in experimental design without fully integrating conceptual understanding, metacognitive regulation, and risk assessment. Importantly, such weaknesses are unlikely to be detected through traditional paper-and-pencil assessments, underscoring the significance of performance-based assessment as an authentic means of capturing complex scientific practices. By requiring pre-service teachers to plan experiments in open-ended contexts, the assessment revealed subtle but consequential gaps that have direct implications for classroom practice.

From a curriculum perspective, the findings are strongly aligned with Ghana's Standards-Based Curriculum for Science, which emphasises inquiry, experimentation, and the development of science process skills as foundations for scientific literacy and lifelong learning. The generally high planning proficiency observed suggests that teacher education programmes are making progress toward these curricular goals. Nevertheless, the recurring neglect of safety and measurement accuracy highlights a disconnect between curriculum intentions and enacted competencies. Given that pre-service teachers will soon be responsible for guiding learners in practical science activities, such gaps pose potential risks to both learner safety and the quality of inquiry-based instruction at the basic school level.

Gender comparisons revealed no substantial or consistent differences in planning skills across most task components, indicating that both male and female pre-service teachers demonstrated comparable levels of proficiency. This finding supports literature suggesting that when learners are exposed to similar instructional opportunities and assessment conditions, gender disparities in science process skills tend to diminish. From an equity and Education for Sustainable Development (ESD) standpoint, this outcome is encouraging, as it reinforces the potential of inquiry-oriented and performance-based approaches to promote inclusive participation and competence development in science education. In sum, the study concludes that while pre-service teachers' exhibit promising levels of procedural planning skills, their planning competence is incomplete without explicit attention to safety reasoning, conceptual accuracy, and reflective decision-making. Addressing these gaps is essential for preparing scientifically literate teachers capable of implementing inquiry-based, learner-centred, and safe science instruction in line with national curriculum expectations and ESD goals.

5.2 Recommendations

It is recommended that, Colleges of Education should explicitly integrate laboratory safety reasoning into the teaching and assessment of science process skills. Safety should be treated as a core component of planning rather than an implicit or peripheral consideration. Lecturers should model safety-conscious planning during demonstrations and require pre-service teachers to justify safety measures as part of every experimental design task. Embedding safety criteria into performance task rubrics would signal its importance and encourage deliberate attention during planning.

Second, teacher education programmes should strengthen the integration of conceptual understanding with procedural planning. While pre-service teachers demonstrated the ability to sequence steps, weaknesses in measurement accuracy and formula application indicate the need for instructional approaches that foreground the rationale behind each planning decision. Inquiry-based pedagogies that prompt learners to explain why specific procedures, instruments, or calculations are necessary may help bridge the gap between procedural fluency and conceptual depth.

Third, assessment practices within Colleges of Education should increasingly adopt performance-based and process-oriented assessments alongside traditional tests. As demonstrated in this study, performance tasks provide richer evidence of learners' planning abilities, decision-making processes, and misconceptions. Such assessments are consistent with Ghana's competency-based curriculum and support the development of transferable skills essential for scientific literacy and sustainable development.

Fourth, professional development opportunities should be provided for science teacher educators to enhance their capacity to design, implement, and score performance-based assessments reliably. Training should focus on rubric development, observation of process skills, and feedback strategies that promote reflective practice among pre-service teachers. This will help ensure consistency, validity, and instructional alignment in assessment practices. Fifth, the findings support continued emphasis on gender-inclusive inquiry-based instruction, as no major gender differences were

observed in planning skills. Teacher educators should sustain learning environments that provide equal access to practical experiences, thereby supporting ESD goals related to equity and inclusion in science education.

Finally, future research should explore longitudinal development of planning skills among pre-service teachers, tracking how these competencies evolve during supported teaching in schools and early career practice. Qualitative studies, including interviews and classroom observations, could further illuminate how pre-service teachers conceptualise planning and safety, and how assessment experiences influence their instructional practices. Expanding the scope to multiple Colleges of Education would also enhance the generalisability of findings. Overall, strengthening the holistic development of planning skills through inquiry-based teaching and performance-based assessment will better prepare pre-service teachers to implement Ghana's science curriculum effectively, promote learner safety, and foster scientific literacy aligned with national and global ESD aspirations.

Declaration of Interest

The authors declare that they do not have any known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Alalouch, C. (2021). Cognitive styles, gender, and student academic performance in engineering education. *Education Sciences, 11*(9), 502.

Antwi, V., Sakyi-Hagan, N. A., Addo-Wuver, F., & Asare, B. (2021). Effect of practical work on physics learning effectiveness: A case of a senior high school in Ghana. *East African Journal of Education and Social Sciences, 2*(3), 43–55.

Atasoy, V., & Kaya, G. (2022). Formative assessment practices in science education: A meta-synthesis study. *Studies in Educational Evaluation, 75*, 101186.

Beroíza-Valenzuela, F. (2025). Implicit gender stereotypes in STEM: Measuring cognitive bias and group differences through reaction times. *International Journal of STEM Education, 12*(1), 20.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice, 25*(6), 551–575.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24*(2), 97–140. https://doi.org/10.1080/10888691.2018.1537791

Dolapcioglu, S., & Subasi, M. (2022). The relationship between scientific process skills and science achievement: A meta-analysis study. *Journal of Science Learning, 5*(2), 363–372.

García-Carmona, A., Criado, A. M., & Cruz-Guzmán, M. (2017). Primary pre-service teachers' skills in planning a guided scientific inquiry. *Research in Science Education, 47*(5), 989–1010.

Goles-Sabellina, J. (2021). Classroom assessment integration and practices of college instructors: Basis for professional training design on assessment. *Journal of Innovations in Teaching and Learning, 1*(1), 47–55.

Hammann, M., Phan, T. T. H., Ehmer, M., & Bayrhuber, H. (2008). Students' difficulty in designing experiments: A review of empirical research. *Educational Research Review, 3*(1), 1–24. https://doi.org/10.1016/j.edurev.2007.06.003

Hyytinen, H., Ursin, J., Silvennoinen, K., Kleemola, K., & Toom, A. (2021). The dynamic relationship between response processes and self-regulation in critical thinking assessments. *Studies in Educational Evaluation, 71*, 101090.

Johnson, E. A. (2017). Assessment of Ghanaian senior high school students' science process skills in optics (Doctoral dissertation, University of Cape Coast).

Jumaah, F. M. (2024). Constructivist approaches and learner engagement in science classrooms. *Journal of Educational Research and Practice, 14*(1), 112–126. https://doi.org/10.5590/JERAP.2024.14.1.09

Kashikar, L., Lüke, T., & Grosche, M. (2026). Effects of diagnostic labels for students with learning problems on teachers' stereotypes and performance expectations. *Journal of Learning Disabilities, 59*(1), 37–54.

Katayev, M. (2023). Science process skills as predictors of inquiry competence in secondary science education. *Research in Science Education, 53*(2), 431–450. https://doi.org/10.1007/s11165-021-10027-6

Koomson, A. (2020). An investigation into development of science process skills by senior high school chemistry students (Doctoral dissertation, University of Cape Coast).

Koomson, A., Kwaah, C. Y., & Adu-Yeboah, C. (2024). Effect of science process skills and entry grades on academic scores of student teachers. *Journal of Turkish Science Education, 21*(1), 118–133.

Levy-Feldman, I. (2025). The role of assessment in improving education and promoting educational equity. *Education Sciences, 15*(2), 224. https://doi.org/10.3390/educsci15020224

National Council for Curriculum and Assessment [NACCA]. (2019). *Standards-based curriculum for basic schools: Science curriculum*. Ministry of Education, Ghana.

Nicol, C. B., Gakuba, E., & Habinshuti, G. (2022). Students' perceived science inquiry process skills in relation to school type and gender. *Perspectives in Education, 40*(2), 159–174.

Ogunsola-Bandele, M., & Kennepohl, D. (2024). Alternatives to traditional physical science laboratory experiences: A review. *West African Journal of Open and Flexible Learning, 12*(2), 165–200.

Organisation for Economic Co-operation and Development [OECD]. (2023). *PISA 2022 assessment and analytical framework: Science, reading, mathematics, and creative thinking*. OECD Publishing. https://doi.org/10.1787/2023.978926

Photo, P. (2025). Teaching experiences connected to the implementation of inquiry-based practical work in primary science classrooms. *Research in Science Education, 55*, 1493–1516. https://doi.org/10.1007/s11165-025-10235-3

Rozgonjuk, D., Täht, K., Soobard, R., Teppo, M., & Rannikmäe, M. (2024). The S in STEM: Gender differences in science anxiety and its relations with science test performance-related variables. *International Journal of STEM Education, 11*(1), 45.

Schoenfeld, A. H. (2023). Commentary on Part III of mathematical challenges for all: On problems, problem-solving, and thinking mathematically. In *Mathematical challenges for all* (pp. 565–584). Cham: Springer International Publishing.

Teig, N. (2024). Uncovering student strategies for solving scientific inquiry tasks: Insights from student process data in PISA. *Research in Science Education, 54*(2), 205–224.

Temiz, B. (2020). Assessing skills of identifying variables and formulating hypotheses using scenario-based multiple-choice questions. *International Journal of Assessment Tools in Education, 7*(1), 1–17.

Vlachopoulos, D., & Makri, A. (2024). A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation, 83*, 101425.

Wren, D. G. (2019). *Assessing deeper learning: Developing, implementing, and scoring performance tasks*. Rowman & Littlefield.

Yaghoubi, M., Pourhosein Gilakjani, A., & Abbasian, G. R. (2025). The effect of performance-based assessment on academic resilience, motivation, and teacher support in virtual versus non-virtual classes. *Language Testing in Asia, 15*(1), 1–32.

Yakob, M., Hamdani, H., Sari, R. P., Haji, A. G., & Nahadi, N. (2021). Implementation of performance assessment in STEM-based science learning to improve students' habits of mind. *International Journal of Evaluation and Research in Education, 10*(2), 624–631.