African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27)
www.abjournals.org

FROM DATA HANDLING TO EVIDENCE-BASED REASONING: LATENT PROFILES OF JUNIOR HIGH SCHOOL LEARNERS' STATISTICAL REASONING IN GHANA

Benjamin Apam

Department of Statistics, Bolgatanga Technical University.

Email: bapam@bolgatu.edu.gh

Cite this article:

Benjamin Apam (2026), From Data Handling to Evidence-Based Reasoning: Latent Profiles of Junior High School Learners' Statistical Reasoning in Ghana. African Journal of Mathematics and Statistics Studies 9(3), 1-27. DOI: 10.52589/AJMSS-P3QSH4EW

Manuscript History

Received: 20 Jun 2026

Accepted: 16 Jul 2026

Published: 28 Jul 2026

Copyright © 2026 The Author(s). This is an Open Access article distributed under the terms of Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), which permits anyone to share, use, reproduce and redistribute in any medium, provided the original author and source are credited.

ABSTRACT: *This study examined how Junior High School learners' statistical reasoning is configured across data representation, computation, interpretation, comparison, and evidence-based conclusion drawing. Drawing on a cross-sectional assessment of 437 JHS 2 learners from 13 schools in Bolgatanga Municipality, Ghana, the study used latent profile analysis to identify learner subgroups and multinomial logistic regression to examine predictors of profile membership. A three-profile solution was retained: emerging statistical reasoning, developing procedural reasoning, and advanced balanced reasoning. The largest profile showed relatively stronger computation and data representation but weaker interpretation, comparison, and evidence-based conclusion drawing, indicating that procedural data handling does not necessarily translate into statistical reasoning with evidence. Prior mathematics achievement, textbook access, private school attendance, data-task exposure, and school location were associated with profile membership. The study contributes to statistics education research by providing a construct-linked, person-centred account of how school-level statistical reasoning varies within an under-represented African basic education context and by showing how profile evidence can inform differentiated instruction in data-rich mathematics classrooms.*

KEYWORDS: Statistical reasoning; Statistical literacy; Latent profile analysis; Mathematics education; Diagnostic assessment.

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

INTRODUCTION

The ability to reason with data is now a central educational competence because learners are expected to interpret graphs, compare numerical information, evaluate claims, recognise variation and uncertainty, and use evidence to support conclusions in school and everyday life (Bargagliotti et al., 2020; Ben-Zvi & Garfield, 2004; Gal, 2002; Garfield & Ben-Zvi, 2008; Watson, 2006; Wild & Pfannkuch, 1999). In statistics education, this competence is usually discussed through the related but distinct constructs of statistical literacy, statistical reasoning, and statistical thinking. Statistical literacy emphasises interpreting, evaluating, and communicating statistical information in context; statistical reasoning concerns how learners make sense of statistical ideas such as data, representation, centre, variability, comparison, distribution, and inference; and statistical thinking extends to the broader investigative cycle of asking questions, producing or obtaining data, analysing patterns, and drawing warranted conclusions under uncertainty (Ben-Zvi & Garfield, 2004; Gal, 2002; Garfield, 2002; Wild & Pfannkuch, 1999).

This distinction matters because school statistics can easily be reduced to routine "data handling": completing a frequency table, drawing a bar chart, finding a mean, or reading off the largest value. These tasks are important, but they are not sufficient indicators of statistical reasoning. A learner may compute accurately while still failing to explain what a result means, compare groups appropriately, recognise variation, or evaluate whether a conclusion is supported by data. The present study therefore treats statistical reasoning as coordinated reasoning with data, context, representation, variation, comparison, and evidence rather than as a single mathematics achievement score.

Empirical findings demonstrate that learners' performance can vary across different components of statistical reasoning. Jones et al. (2000) examined 20 learners in Grades 1-5 and found that 80% demonstrated a stable level of thinking across at least three of four processes: describing, organising, representing, and analysing and interpreting data. Mooney's (2002) study of 12 learners in Grades 6-8 similarly identified four developmental levels within these processes, showing that the sophistication of learners' reasoning could differ across statistical tasks. More recent evidence has reinforced this pattern. Balkaya and Kurt (2023) found that middle school learners' reasoning about distributions developed through creating, reviewing, comparing, and making conjectures about statistical models rather than through producing displays alone. In a longitudinal study of 44 learners, Oslington et al. (2024) found that predictions within the historical data range increased from 54% in Year 3 to 87% in Year 4, while data-informed explanations increased from 50% to 79%. By Year 4, 57% of learners also produced representations demonstrating meaningful transformation of the original data. Nevertheless, representations and explanations generally remained less developed than prediction accuracy, indicating uneven progress across the assessed reasoning processes.

The relevance of this evidence to Ghanaian Junior High School (JHS) mathematics lies in the absence of comparable empirical knowledge about how these reasoning processes are configured within individual learners. The Common Core Programme requires learners to engage with data handling, frequency tables, graphical representations, measures of central tendency, problem-solving, and the interpretation of numerical information (National Council for Curriculum and Assessment [NaCCA], n.d.). However, existing classroom and examination summaries generally report total mathematics marks or broad achievement categories and therefore provide limited evidence about whether learners' strengths and difficulties are

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27) www.abjournals.org

concentrated in representation, computation, interpretation, comparison, or evidence-based conclusion drawing. Moreover, much of the established research has been conducted outside African basic education systems, where language conditions, instructional resources, classroom practices, and assessment environments may differ from those experienced by Ghanaian learners (Balkaya & Kurt, 2023; Garfield & Ben-Zvi, 2008; Oslington et al., 2024; Watson, 2006). A person-centred approach is therefore appropriate because it can identify subgroups of learners with similar configurations across multiple reasoning indicators rather than relying solely on total scores or average relationships (Collins & Lanza, 2010; Masyn, 2013; Nylund-Gibson & Choi, 2018). Latent profile analysis is principally suitable where the indicators are continuous scores in data representation, computation, interpretation, comparison, and evidence-based conclusion drawing (Berlin et al., 2014; Spurk et al., 2020). Accordingly, the present study examined whether JHS 2 learners in Bolgatanga Municipality formed distinct statistical reasoning profiles, which learner-level and school-level characteristics predicted profile membership, and whether the profiles differed in mathematics confidence, classroom engagement, and prior mathematics achievement.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Statistical reasoning is coordinated reasoning with data, context and evidence

A useful theoretical starting point is the view that statistical reasoning develops through learners' increasingly sophisticated ways of organising, reducing, representing, analysing, interpreting, and using data. Jones et al. (2000) and Mooney (2002) characterised children's and middle school students' statistical thinking through processes such as describing data, organising and reducing data, representing data, and analysing and interpreting data. These frameworks are relevant to the present study because they show that school-level statistical reasoning is not a single skill but a set of related processes that may develop unevenly across learners.

The statistics education literature also emphasises that reasoning with data requires attention to aggregate patterns rather than isolated values. Learners often begin by focusing on individual cases or visible features of displays before gradually coordinating data values, graphical structure, centre, spread, variability, and contextual meaning (Ben-Zvi & Garfield, 2004; Garfield & Ben-Zvi, 2008; Konold & Higgins, 2003; Lehrer & Schauble, 2004; Watson, 2006). Recent studies continue to show that students' representations, interpretations, and explanations do not always develop at the same pace. For example, research on predictive reasoning from data tables found that learners could make more developed predictions over time, but their representations and explanations could still lag behind their predictions (Oslington et al., 2024). Similarly, work on middle school students' reasoning about distribution shows the importance of creating, reviewing, comparing, and conjecturing about data models rather than merely producing displays (Balkaya & Kurt, 2023).

Informal statistical inference adds another important layer to this framework. Makar and Rubin (2009) argue that informal inference involves making generalisations beyond the immediate data, using data as evidence, and expressing uncertainty in non-deterministic ways. For JHS learners, this does not require formal sampling distributions or hypothesis tests. It means being able to judge whether a conclusion is reasonable, whether a claim exaggerates the evidence,

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

whether two groups should be compared by raw counts or rates, and whether similar averages necessarily imply identical data. Evidence-based conclusion drawing in the present study was therefore conceptualised as an early form of data-based argumentation: learners had to connect a claim to relevant evidence and explain why the claim was or was not supported.

Empirical basis for the five assessment domains

The five domains used in this study were derived from this theoretical and empirical literature. Data representation captures the organisation and display of data, including the movement from data values to tables, graphs, and meaningful descriptions. Computation captures the reducing and summarising work needed to produce frequencies, totals, centre, or ranges. Interpretation captures learners' ability to explain what a display, pattern, or summary means in context. Comparison captures the ability to compare groups, rates, change, stability, and variability using appropriate evidence. Evidence-based conclusion drawing captures the ability to use data to support or challenge claims. Together, the domains provide a school-level operationalisation of statistical reasoning that is broad enough to capture procedural, representational, interpretive, comparative, and evidential components, but be specific enough to be scored and modelled as profile indicators.

The distinction among these domains is theoretically important for a person-centred study. If all learners simply differed by one general level of achievement, a total score would be sufficient. However, prior research suggests that learners can be strong in one aspect of statistical work and weaker in another; for example, they may construct displays but struggle to reason about variation, or they may compute correctly but fail to justify a conclusion. This makes latent profile analysis appropriate because it can identify configurations of strengths and weaknesses across multiple reasoning indicators instead of collapsing them into a single score.

The Ghanaian basic education context and the need for diagnostic evidence

In Ghanaian Junior High School mathematics, learners encounter data handling, frequency tables, graphs, measures of central tendency, and interpretation of information. The Common Core Programme places JHS within Key Phase 4 and emphasises core competencies, learner-centred learning, problem-solving, and preparation for further study and adult life (National Council for Curriculum and Assessment [NaCCA], n.d.). These curricular ambitions align with international calls for statistics learning that connects data, context, variation, and interpretation. Yet assessment reports and classroom summaries often reduce mathematics performance to total marks or broad achievement bands. Such summaries may indicate whether learners are performing well or poorly, but they do not show whether learners' difficulties lie in representation, computation, interpretation, comparison, or evidence-based conclusion drawing.

The Ghanaian context also matters because learners' everyday lives contain rich data contexts, including school attendance, rainfall, market prices, household expenditure, farming activities, public health information, and sports results. These contexts can support meaningful statistics learning when classroom tasks deliberately connect local data to interpretation and evidence. At the same time, learners' reasoning may be shaped by language of instruction, access to textbooks, school location, classroom resources, and opportunities to practise data-related tasks. For an international SERJ readership, the value of the Ghanaian case is not that it is simply a new location but that it offers evidence about how diagnostic statistical-reasoning

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27)
www.abjournals.org

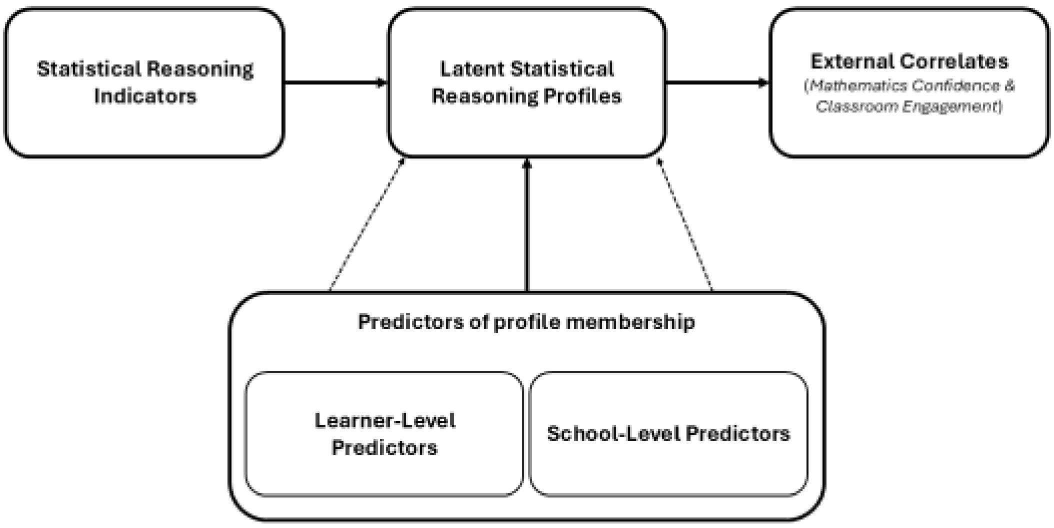
profiles may appear in a basic education system where data-rich everyday contexts coexist with resource and instructional constraints.

Person-centred modelling as a diagnostic lens

Person-centred methods are useful when learner populations are heterogeneous and when average scores conceal educationally meaningful subgroups. Latent profile analysis identifies unobserved profiles of learners who share similar patterns across continuous indicators (Collins & Lanza, 2010; Masyn, 2013; Nylund-Gibson & Choi, 2018; Spurk et al., 2020). In the present study, the indicators were the five statistical-reasoning domain scores. The aim was not to label learners permanently but to identify interpretable profiles that could support differentiated instruction. A profile showing uniformly low performance would call for broad foundational support; a profile showing stronger computation but weaker interpretation and evidence use would call for tasks that connect procedures to meaning, comparison, and justification; and a balanced high-performing profile would suggest readiness for more complex data investigations and informal inference.

Fig. 1 summarises the conceptual relationships among the statistical reasoning indicators, learner-level and school-level characteristics, latent profile membership, and the learning-related variables examined across the identified profiles.

Fig. 1: Conceptual framework of statistical reasoning profile membership and related variables



Purpose, research questions and contribution

The study therefore asks three research questions:

(1) What distinct latent profiles of statistical reasoning can be identified among JHS learners in Ghana?

(2) Which learner- and school-level characteristics predict membership in the identified profiles?

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

(3) How do mathematics confidence, classroom engagement, and prior mathematics achievement differ across the identified profiles?

The contribution of the study is threefold. First, it provides an empirically grounded profile account of school-level statistical reasoning in an under-represented African basic education context. Second, it links the measurement of statistical reasoning explicitly to theory by modelling five construct-derived domains rather than a single total score. Third, it shows how profile evidence can support diagnostic and differentiated mathematics instruction by distinguishing procedural data handling from evidence-based statistical reasoning.

METHODS

Research design and setting

The study used a cross-sectional quantitative design to identify latent profiles of statistical reasoning among JHS learners in the Bolgatanga Municipality in the Upper East Region of Ghana. Cross-sectional data are commonly used in latent profile analysis when the objective is to identify unobserved subgroups of individuals with similar response patterns across continuous indicators (Berlin et al., 2014; Collins & Lanza, 2010; Masyn, 2013; Nylund-Gibson & Choi, 2018; Spurk et al., 2020).

The study was framed as a diagnostic investigation of mathematics learning and not a validation study, because its central concern was whether learners' performance across data representation, computation, interpretation, comparison, and evidence-based conclusion drawing could provide evidence for more targeted instruction. Several authors alluded to the fact that diagnostic assessment is useful when it reveals how learners understand, misunderstand, or partially understand a concept and when such evidence informs instructional decisions (Black & Wiliam, 1998; Heritage, 2010; Pellegrino et al., 2001). The conceptualisation of statistical reasoning as a multidimensional construct was consistent with literature on statistical literacy, statistical reasoning, and statistical thinking (Bargagliotti et al., 2020; Ben-Zvi & Garfield, 2004; Franklin et al., 2007; Gal, 2002; Garfield, 2002; Garfield & Ben-Zvi, 2008; Watson & Callingham, 2003; Wild & Pfannkuch, 1999).

Although data were collected at one time point, the analysis recognised that learners were nested within classrooms and schools. Learners in the same school may share instructional exposure, teacher practices, peer learning conditions, school resources, and assessment culture. Ignoring such clustering may underestimate standard errors and overstate statistical precision in school-based educational research (Hox et al., 2017; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). The analysis therefore began with conventional latent profile models and assessed whether design-based corrections, cluster-robust standard errors, or multilevel mixture modelling were necessary.

Bolgatanga Municipality was suitable for the study because it is the administrative capital of the Upper East Region and includes schools serving learners from varied socio-economic, linguistic, and settlement backgrounds. The Ghana Statistical Service (GSS) reported the municipality's 2021 population as 139,864, comprising 66,607 males and 73,257 females, and representing 10.7% of the regional population. The Municipal Assembly describes the municipality as bordered by Bongo District to the north; Talensi and Nabdam Districts to the

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

south and east; and Kassena-Nankana East and West Districts to the west. The municipality also provides meaningful everyday data contexts, including market prices, rainfall variation, farming activities, school attendance, household expenditure, sports results, public health messages, and demographic information. Such contexts are relevant because statistics education research emphasises that learners develop stronger statistical reasoning when data are connected to meaningful situations and when classroom tasks require interpretation, comparison, and evidence-based judgement rather than routine calculation alone (Bargagliotti et al., 2020; Cobb & McClain, 2004; Franklin et al., 2007; Garfield & Ben-Zvi, 2008; Makar & Rubin, 2009; Pfannkuch, 2011).

Population, sampling and sample size

The target population comprised JHS learners enrolled in basic schools in Bolgatanga Municipality. The accessible population consisted of learners in selected public and private JHS within the municipality during the period of data collection. The study focused on JHS 2 learners because they were expected to have encountered foundational data-handling concepts while not yet being under the full examination pressure associated with JHS 3. Focusing on JHS 2 also reduced curriculum heterogeneity. Participation was not restricted by prior achievement level because the aim was to identify the full range of statistical reasoning profiles present in ordinary classrooms. A multistage sampling procedure was used. First, a list of JHS in Bolgatanga Municipality was obtained from the Municipal Education Directorate. Second, schools were stratified by school ownership and school location. School ownership distinguished public and private schools, while school location distinguished schools in more urbanised parts of the municipality from those in peri-urban or rural communities.

Third, schools were selected from each stratum using proportionate or approximately proportionate sampling, depending on the distribution of schools and fieldwork feasibility. Fourth, intact JHS 2 classrooms in the selected schools were sampled, and all eligible learners in those classrooms were invited to participate. The use of intact classrooms was practical because it limited disruption to school schedules and reflected actual instructional units. Because intact classroom sampling creates clustered data, the analysis accounted for classroom and school clustering where necessary (Hox et al., 2017; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Drawing learners from multiple schools also made it possible to examine whether profile membership was associated with school ownership, school location, class size, and resource context.

The sample size was determined using Cochran’s formula (Cochran, 1977) because the target population was large, the proportion of learners exhibiting different levels of statistical reasoning was unknown, and a minimum sample was required at a 95% confidence level.

$$n_0 = \frac{z^2 p(1-p)}{e^2}$$

where n_0 is the required initial sample size, z is the standard normal value corresponding to the selected confidence level; p is the estimated proportion of the population with the attribute of interest, and e is the desired margin of error. Because no prior empirical estimate was available on the proportion of JHS learners in Bolgatanga Municipality who belonged to specific statistical reasoning profiles, $p = 0.50$ was used. This conservative choice maximises

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

the required sample size and provides adequate precision when the true population proportion is unknown (Bartlett et al., 2001; Cochran, 1977; Israel, 1992).

Using a 95% confidence level, $z = 1.96$; a 5% margin of error, $e = 0.05$; and $p = 0.50$, the sample size was calculated as follows:

$$n_0 = \frac{z^2 p(1-p)}{e^2} = \frac{1.96^2 \times 0.50 \times (1-0.50)}{0.05^2} = 384.16 \approx 385$$

The minimum required sample size was therefore approximately 385 learners. To account for possible non-response, incomplete scripts, absenteeism, and unusable questionnaires, the sample size was increased by 10%. The adjusted sample size was calculated as:

$$n_{adj} = \frac{n_0}{1-r} = \frac{385}{1-0.10} = 427.78 \approx 428$$

where n_{adj} is the adjusted sample size, n_0 is the initial sample size, and r is the anticipated non-response rate.

For field implementation, 13 JHS were selected, with one intact JHS 2 classroom drawn from each school. The selected classrooms had an average class size of 35 learners, giving a field sample of 455 learners before exclusions due to absenteeism, non-response, or incomplete scripts. After data screening, 437 valid learner responses were retained for analysis. This final analytic sample exceeded the adjusted minimum sample size and represented approximately 96.0% of the projected field sample.

Measures

The main measure was a Statistical Reasoning Assessment (SRA) developed for JHS learners in Bolgatanga Municipality. The SRA was designed from the theoretical position that school-level statistical reasoning involves coordinated work across representation, reduction/computation, interpretation, comparison, and data-based justification. The instrument therefore covered five domains: data representation, computation, interpretation, comparison, and evidence-based conclusion drawing. Each domain contained four items scored from 0 to 3, yielding a maximum of 12 per domain and 60 for the full assessment. The five domain scores, not the total score, were used as the latent profile indicators because the purpose was to identify patterns of reasoning across domains.

Data representation assessed learners' ability to organise, display, and describe data using grade-appropriate formats such as tables, bar charts, pictographs, and simple line graphs. Computation assessed the ability to reduce and summarise data through frequencies, totals, means, modes, and range. Interpretation assessed the ability to explain what a table, graph, pattern, or summary measure meant in context. Comparison assessed the ability to compare categories, groups, rates, change, or stability and make defensible statements about similarities and differences. Evidence-based conclusion drawing assessed the ability to use data to support a claim, reject an unsupported claim, or explain why a conclusion followed from the available evidence.

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

The assessment included selected-response and constructed-response items as mapped in Table 1. Selected-response items were used where the reasoning demand involved recognising suitable representations, reading displays, or identifying correct summaries. Constructed-response items were used where explanation, comparison, or justification was required because learners' statistical reasoning cannot be fully inferred from final answers alone (Ben-Zvi & Garfield, 2004; Garfield, 2002; Garfield & Ben-Zvi, 2008; Makar & Rubin, 2009; Watson, 2006). Constructed responses were scored with an analytic rubric distinguishing no evidence of reasoning, limited reasoning, partially correct reasoning, and well-supported statistical reasoning. The full assessment and scoring guide are provided in the Supplementary File.

Table 1. Construct map linking theoretical reasoning domains to assessment evidence

Domain	Theoretical reasoning demands	Assessment tasks	Scoring evidence	LPA indicator
Data representation	Organising and representing data: transnumeration from values to display	C01-C04: selecting, drawing, completing and describing displays	Appropriate display choice, accurate structure, labels, values and descriptive statement	Domain score 0-12
Computation	Reducing data through numerical summaries and basic measures	C05-C08: mean, total, mode and range	Correct method, accurate calculation and interpretable working where required	Domain score 0-12
Interpretation	Reading and explaining data in context, including patterns and non-equivalence of equal summaries	C09-C12: graph/table reading, trend interpretation and meaning of equal averages	Contextual explanation linked directly to values, patterns or statistical meaning	Domain score 0-12
Comparison	Comparing groups, rates, change, stability and variability	C13-C16: pass rates, change over time and stability of values	Appropriate comparison basis, relevant calculation or evidence, and justified statement	Domain score 0-12

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

Domain	Theoretical reasoning demands	Assessment tasks	Scoring evidence	LPA indicator
Evidence-based conclusion	Using data as evidence to support, qualify or reject claims	C17-C20: judging claims about improvement, evenness, change and magnitude	Conclusion supported by relevant data and explanation of why the claim is reasonable or unreasonable	Domain score 0-12

Source: *Author's construct developed based on the theoretical and empirical literature review.*

Mathematics confidence was measured using a short learner questionnaire on perceived ability to understand mathematics, solve mathematics problems, interpret data displays, and explain mathematical thinking. Classroom engagement was measured with items on attention, participation, task completion, persistence, effort to understand explanations, and willingness to explain reasoning. Data-task exposure captured how often learners encountered tables, graphs, averages, comparisons, real-life examples, and evidence-based questions in mathematics lessons. Prior mathematics achievement was measured using learners' most recent school mathematics scores or a short-standardised mathematics test where school records were unavailable or not comparable across schools. Learner-level variables included sex, age, home language, language used for learning support at home, and availability of learning support. School-level variables included school ownership, school location, class size, mathematics teaching resources, and teacher-reported emphasis on data handling.

Validity and reliability

Several procedures were used to strengthen measurement quality and to make the connection between construct, item demand, scoring evidence, and profile indicators explicit. First, the assessment blueprint was reviewed against the five-domain construct map to confirm that each item elicited the intended form of statistical reasoning rather than merely general reading ability or routine arithmetic. Second, experts in mathematics education, statistics education, and JHS teaching reviewed the tasks for content relevance, grade appropriateness, language clarity, contextual fairness, and alignment with the intended reasoning domain. This step was necessary because assessment tasks must support valid interpretation of learners' understanding rather than merely produce scores (American Educational Research Association et al., 2014; Pellegrino et al., 2001).

Second, the assessment was piloted with JHS learners in a school that was not included in the main study. The pilot examined item clarity, completion time, difficulty level, response distribution, and possible ambiguity in constructed-response scoring. Items showing extreme difficulty, poor discrimination, unclear wording, excessive reading burden, or weak alignment with the intended domain were revised or removed because statistical reasoning assessments can unintentionally measure reading difficulty, contextual unfamiliarity, or test-taking skill (Garfield & Ben-Zvi, 2008; Watson, 2006).

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

Third, constructed-response items were scored using the analytic rubric in the Supplementary File. Raters were trained using sample responses representing different score levels, and at least 20% of scripts were independently scored by two raters. Where agreement was below an acceptable level, raters revisited the scoring guide, discussed disagreements, and rescored affected responses before final data entry.

Fourth, internal consistency was assessed for the full assessment and each domain score using reliability coefficients appropriate to the item structure. Cronbach's alpha was reported where useful, but the study did not rely on alpha alone because alpha has limitations where multidimensionality, unequal item loadings, or heterogeneous item formats are present (McDonald, 1999; Revelle & Condon, 2019; Sijtsma, 2009). Omega coefficients were considered where the domain structure and number of items supported their estimation (Dunn et al., 2014; McDonald, 1999; Revelle & Condon, 2019).

Fifth, the dimensional structure of the assessment was examined before profile modelling using confirmatory factor analysis or exploratory structural equation modelling where the sample size and item structure permitted. This was treated as a measurement check but not a separate validation study because latent profile analysis depends on the quality and interpretability of the indicators used to define profiles (Masyn, 2013; Nylund-Gibson & Choi, 2018; Spurk et al., 2020). Final domain scores were retained after considering expert judgement, pilot evidence, scoring reliability, internal consistency, and empirical distinctiveness.

Data collection and ethical procedures

Permission was obtained from the Bolgatanga Municipal Education Directorate, selected school heads, and classroom teachers before data collection. Data collection was scheduled with participating schools to minimise disruption to teaching and assessment activities.

The statistical reasoning assessment was administered under standardised classroom conditions by trained research assistants. Learners were informed that the assessment was for research purposes only and would not affect their school grades, reducing the risk that high-stakes framing would alter learner behaviour or increase anxiety. The learner questionnaire was administered after the assessment or on the same day, depending on school schedules. Where prior mathematics achievement was obtained from school records, the information was matched to learner responses using anonymous study identification codes.

The learner-facing instrument contained only the statistical reasoning assessment and learner questionnaire. The scoring rubric was not included in the learner-facing instrument and was used only after field administration by trained raters. Research assistants were trained on administration procedures, ethical conduct, learner assent, handling learner questions without coaching, and checking completed instruments. They clarified only general instructions and did not explain how to solve assessment items. Completed assessments were reviewed for missing pages, duplicate identifiers, missing learner codes, and obvious administration problems before data entry. Constructed-response items were then prepared for scoring using the analytic rubric, alongside the survey instrument in the Supplementary File, and data were double-entered or checked using programmed validation rules to reduce entry errors.

The study observed ethical principles for research involving minors in school settings. Participation was voluntary, and learners could decline participation or withdraw without

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

penalty. Parental or guardian consent and learner assent were obtained in line with institutional and education authority requirements. Learner names did not appear in the analysis dataset. Each participant was assigned a unique study code, and school identifiers were anonymised in reporting unless explicit permission was granted and identification was necessary. Profile labels were not used to stigmatise learners, teachers, or schools; rather, latent profiles were interpreted as current patterns of statistical reasoning under the assessment conditions rather than fixed categories of learner ability.

Data analysis

Data analysis was conducted in Stata 15.0 to identify latent statistical reasoning profiles, examine predictors of profile membership, and compare selected distal outcomes across the retained profiles. The dataset was first screened for missing values, outliers, distributional patterns, coding errors, and inconsistent responses. Descriptive statistics were computed for learner characteristics, school-level characteristics, mathematics confidence, classroom engagement, and the five statistical reasoning domain scores. Before profile estimation, constructed-response items were scored using the analytic rubric in the Supplementary File, inter-rater reliability was assessed for a subset of independently scored scripts, and domain-level reliability was examined using alpha and omega coefficients where appropriate (Dunn et al., 2014; McDonald, 1999; Revelle & Condon, 2019; Sijtsma, 2009).

Latent profile analysis was conducted using the five statistical reasoning domain scores as continuous indicators. Let y_i represent the vector of domain scores for learner i. The observed score vector was expressed as:

$$y_i = \left(\text{representation}_i, \text{computation}_i, \text{interpretation}_i, \text{comparison}_i, \text{conclusion}_i\right)$$

The latent profile model assumes that each learner belongs probabilistically to one of K unobserved profiles. The density of learner i's response vector was expressed as a finite mixture model:

$$f(y_i) = \sum_{k=1}^{k} \pi_k f_k(y_i \mid \mu_k, \Sigma_k)$$

where π_k is the proportion of learners in profile k, f_k is the profile-specific multivariate normal density, μ_k is the vector of profile-specific means, and Σ_k is the profile-specific covariance structure. The profile proportions were constrained to sum to one:

$$\sum_{k=1}^{K} \pi_k = 1$$

Models with increasing numbers of profiles were estimated, beginning with a one-profile model and continuing until additional profiles no longer produced a statistically stable, parsimonious, and educationally interpretable solution. For each learner, posterior probabilities of profile membership were computed as:

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

$$P(C_i = k \mid y_i) = \frac{\pi_k f_k(y_i \mid \mu_k, \Sigma_k)}{\sum_{h=1}^{K} \pi_h f_h(y_i \mid \mu_h, \Sigma_h)}$$

where $P(C_i = k \mid y_i)$ is the posterior probability that learner i belongs to profile k, given the learner's observed domain-score pattern. Learners were assigned to the profile with the highest posterior probability for descriptive reporting, while classification quality was evaluated using entropy and average posterior probabilities.

Predictors of profile membership were examined using multinomial logistic regression with the emerging statistical reasoning profile treated as the reference category. For learner i, the model compared the log-odds of membership in profile k relative to the reference profile as:

$$\log\left[\frac{P(C_i = k)}{P(C_i = reference)}\right] = \beta_{0k} + \beta_{1k}X_{1i} + \beta_{2k}X_{2i} + \cdots + \beta_{pk}X_{pi}$$

where C_i is the assigned profile membership for learner i, X_{1i} to X_{pi} are learner-level and school-level predictors, and β_{0k} to β_{pk} are profile-specific regression coefficients. Predictors included sex, age, home language, learning support, prior mathematics achievement, school ownership, school location, class size, and mathematics resource availability. Results were reported as odds ratios, computed as:

$$OR = \exp(\beta)$$

Because profile membership is probabilistic, classification uncertainty was considered in the interpretation of predictor models. The analysis used modal profile assignment for the main multinomial regression and conducted sensitivity checks using posterior probabilities. This approach was reported transparently because classify-analyse procedures may underestimate uncertainty when profile separation is weak (Asparouhov & Muthén, 2014; Bakk & Vermunt, 2016; Nylund-Gibson et al., 2019).

Distal outcomes were compared across the retained profiles to assess their educational relevance. The distal outcomes were mathematics confidence, classroom engagement and prior mathematics achievement. Profile differences were examined using regression-based comparisons with robust standard errors where appropriate. The general comparison model was expressed as:

$$Y_i = \alpha_0 + \alpha_1 \text{Profile}_{2i} + \alpha_2 \text{Profile}_{3i} + \varepsilon_i$$

where Y_i represents the distal outcome for learner i, while the profile indicators compare the developing and advanced profiles with the reference profile. Adjusted means, confidence intervals, *p-values,* and effect sizes were reported. Effect sizes for profile mean differences were computed using standardised mean differences:

$$d = \frac{M_1 - M_2}{SD_{pooled}}$$

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27)

where M_1 and M_2 are the profile-specific means, and $SDpooled$ is the pooled standard deviation. The pooled standard deviation was computed as:

$$SD_{pooled} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}$$

where n_1 and n_2 are the profile-specific sample sizes, and SD_1 and SD_2 are the profile-specific standard deviations. Interpretation emphasised substantive and instructional relevance rather than statistical significance alone.

Methodological limitations

The cross-sectional design allowed the study to identify statistical reasoning profiles but did not support causal conclusions about how those profiles developed. Associations between learner characteristics, school context, and profile membership were therefore interpreted as correlational. The study was also geographically limited to Bolgatanga Municipality, meaning the results should not be generalised uncritically to all JHS learners in Ghana. However, the localised design was appropriate because the study aimed to generate context-sensitive diagnostic evidence for mathematics instruction in a specific municipal education setting.

The study also depended on the quality of the statistical reasoning assessment and the adequacy of domain scores as indicators of learners' reasoning. Although expert review, pilot testing, rubric-guided scoring, inter-rater reliability assessment, internal consistency analysis, and dimensional checks were used to strengthen measurement quality, no single assessment could capture the full complexity of learners' statistical reasoning. Finally, school-level predictors were interpreted cautiously because multilevel estimates may be unstable with too few school clusters.

RESULTS AND DISCUSSION

Results

Table 2 describes the analytic sample of 437 JHS 2 learners. Female learners formed a slight majority of the sample (54.7%), while males accounted for 45.3%. Most learners were aged 13-15 years (80.1%), which is consistent with the study's focus on JHS 2 learners. A smaller proportion were aged 16-18 years (16.5%), while learners aged 10-12 years and 19 years or above represented only 2.3% and 1.1%, respectively.

The home-language distribution reflected the multilingual context of Bolgatanga Municipality. Gurune/Frafra was the dominant home language (63.4%), followed by Hausa (11.2%), Kusaal (10.8%), other languages (8.9%), English (3.0%), and Twi (2.7%). This distribution is important for interpreting statistical reasoning because learners were assessed in a school mathematics context where English is the formal medium of instruction, while everyday reasoning and explanation may be shaped by local language use. Since statistical reasoning requires learners to interpret, compare, and justify conclusions from data, language background provides useful contextual information for understanding learners' performance patterns.

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

Although 63.8% of learners reported having a mathematics textbook at home, 36.2% did not. Digital access was more limited, as only 16.2% reported often having access to a smartphone, tablet, or computer for learning, while 47.8% had access sometimes and 35.9% had no access. These results suggest that many learners may have limited opportunities for independent practice or technology-supported data exploration outside school.

The results show that 300 learners (68.6%) were drawn from public schools and 137 learners (31.4%) from private schools. Similarly, 174 learners (39.8%) were drawn from urban schools, 132 learners (30.2%) from rural schools, and 131 learners (30.0%) from peri-urban schools. This distribution shows that the sample captured learners from different school contexts within the municipality.

Table 2: Sample characteristics of learners and schools (N = 437)

Characteristics	**Category**	***n***	**%**
Sex	Female	239	54.7
	Male	198	45.3
Age group	13-15	350	80.1
	16-18	72	16.5
	10-12	10	2.3
	19+	5	1.1
Home language	Gurune/Frafra	277	63.4
	Hausa	49	11.2
	Kusaal	47	10.8
	Other	39	8.9
	English	13	3.0
	Twi	12	2.7
Mathematics textbook at home	Yes	279	63.8
	No	158	36.2
Digital device for learning	Sometimes	209	47.8
	No	157	35.9
	Often	71	16.2
School ownership	Public	300	68.6
	Private	137	31.4
School location	Urban	174	39.8
	Rural	132	30.2
	Peri-urban	131	30.0

Source: *Field data, 2026*

Table 3 presents the descriptive statistics for the five statistical reasoning domains and selected learning outcomes. Learners performed highest in computation ($M = 7.62$, $SD = 2.20$), followed by data representation ($M = 6.88$, $SD = 2.42$). Lower mean scores were observed for comparison ($M = 5.88$, $SD = 2.29$), interpretation ($M = 5.77$, $SD = 2.29$), and evidence-based conclusion drawing ($M = 5.34$, $SD = 2.22$). Since each domain was scored out of 12, the pattern suggests that learners were relatively stronger in procedural and representational aspects of statistical reasoning than in the more interpretive and evidential domains.

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27) www.abjournals.org

The total statistical reasoning score averaged 31.48 out of 60 (*SD* = 9.21), indicating moderate overall performance with substantial variation among learners. Mathematics confidence (*M* = 3.17, *SD* = 0.46), classroom engagement (*M* = 3.24, *SD* = 0.36) and data-task exposure (*M* = 2.95, *SD* = 0.37) were also moderate, while prior mathematics achievement showed wider variation (*M* = 55.86, *SD* = 10.43).

Table 3: Descriptive statistics for statistical reasoning domains and key learning outcomes

Variable	Mean	SD	Min	Max
Data representation	6.88	2.42	1.00	12.00
Computation	7.62	2.20	1.33	12.00
Interpretation	5.77	2.29	0.00	12.00
Comparison	5.88	2.29	0.00	12.00
Evidence-based conclusion	5.34	2.22	0.00	12.00
Total statistical reasoning	31.48	9.21	13.00	55.33
Mathematics confidence	3.17	0.46	2.00	4.33
Classroom engagement	3.24	0.36	2.17	4.33
Data-task exposure	2.95	0.37	1.71	4.14
Prior mathematics score	55.86	10.43	30.00	84.00

Source: *Author's analysis of field data, 2026*

Table 4 presents the latent profile enumeration statistics used to determine the optimal number of statistical reasoning profiles. Model fit improved substantially from the one-profile to the three-profile solution, as shown by reductions in information criteria and improved classification quality. The three-profile model was retained because it offered the strongest balance of BIC, SABIC, entropy, average posterior probability, profile size, parsimony, and substantive interpretability. Although AIC continued to decline slightly in the four- and five-profile models, the additional profiles reduced classification clarity and did not add sufficient educational interpretability. This supports a parsimonious three-profile solution, consistent with guidance that profile enumeration should consider statistical fit, classification quality, profile size, theoretical meaning, and practical usefulness rather than relying on a single index (Berlin et al., 2014; Masyn, 2013; Nylund et al., 2007; Nylund-Gibson & Choi, 2018; Spurk et al., 2020).

Table 4: Latent profile enumeration statistics

ID	Log likelihood	Parameters	AIC	Entropy	Avg. posterior probability	Smallest profile n	Smallest profile %	BIC	SABIC
1	-3100.38	10	6220.76		1.000	437	100.0	6261.56	6229.82
2	-2723.21	21	5488.41	0.816	0.943	157	35.9	5574.10	5507.46
3	-2574.21	32	5212.42	0.851	0.933	88	20.1	5342.98	5241.43
4	-2560.21	43	5206.42	0.730	0.838	85	19.5	5381.86	5245.40
5	-2546.44	54	5200.87	0.753	0.840	66	15.1	5421.20	5249.83

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

ID	Log likelihood	Parameters	AIC	Entropy	Avg. posterior probability	Smallest profile n	Smallest profile %	BIC	SABIC
6	-2536.56	65	5203.11	0.764	0.825	49	11.2	5468.32	5262.04

Source: *Author's analysis of field data, 2026*

As shown in Table 5, the retained three-profile solution produced clearly differentiated groups of learners. The largest group was the developing procedural reasoning profile (n = 215, 49.20%), followed by the emerging statistical reasoning profile (n = 134, 30.66%) and the advanced balanced reasoning profile (n = 88, 20.14%). Mean posterior probabilities were high across the three profiles, ranging from 0.923 to 0.962, indicating good classification quality. The emerging profile recorded the lowest total reasoning score (M = 21.30) with particularly weak performance in comparison (M = 3.86) and evidence-based conclusion drawing (M = 3.35). The developing procedural profile had a higher total reasoning score (M = 32.05) and showed stronger performance in computation (M = 8.22) and data representation (M = 7.13) than in interpretation (M = 5.52), comparison (M = 5.85), and evidence-based conclusion drawing (M = 5.33). The advanced balanced profile recorded the highest total reasoning score (M = 45.60) and showed consistently high performance across the five domains, although evidence-based conclusion drawing (M = 8.39) remained the relatively lowest domain even within this strongest group.

Fig. 2 further illustrates the shape of the retained profiles across the five statistical reasoning domains. The profile pattern shows that learners differed not only in overall level of statistical reasoning but also in the structure of their reasoning across domains. The developing procedural profile is especially important because it shows that nearly half of the learners performed comparatively better on computational and representational tasks but were less strong in explaining, comparing, and justifying conclusions from data. This distinction supports the study's argument that statistical reasoning should not be inferred from total scores or procedural performance alone. In instructional terms, learners in the emerging profile require foundational support across all reasoning domains, learners in the developing procedural profile require tasks that move them from calculation and representation towards interpretation and evidence-based justification, and learners in the advanced balanced profile may benefit from richer data investigations and informal inference tasks.

Table 5: Retained latent profile sizes and mean domain scores

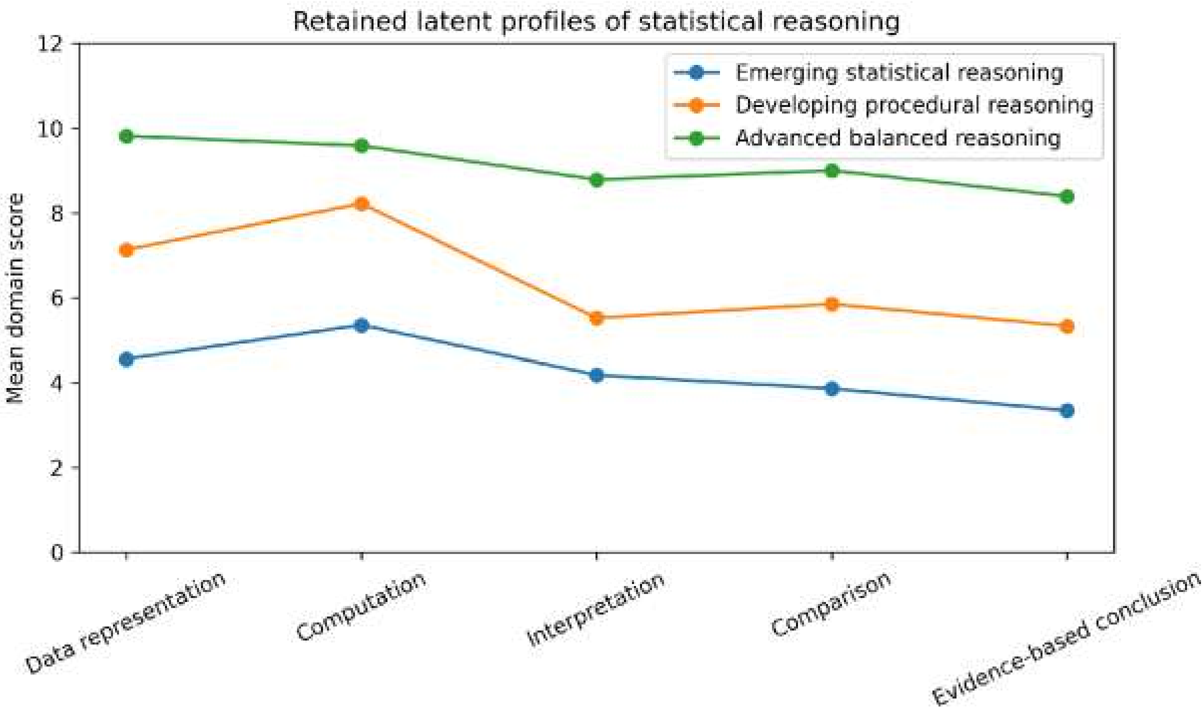
Profile	n	%	Data representation	Computation	Interpretation	Comparison	Evidence-based conclusion	Total reasoning	Mean posterior probability
Emerging statistical reasoning	134	30.66	4.55	5.35	4.18	3.86	3.35	21.30	0.923
Developing procedural reasoning	215	49.20	7.13	8.22	5.52	5.85	5.33	32.05	0.928
Advanced balanced reasoning	88	20.14	9.82	9.59	8.79	9.00	8.39	45.60	0.962

Source: *Author's analysis of field data, 2026*

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27)

Fig. 2. Retained latent profiles of statistical reasoning



Source: *Author's analysis of field data, 2026*

Table 6 presents the multinomial logistic regression results predicting membership in the developing procedural reasoning and advanced balanced reasoning profiles, with the emerging statistical reasoning profile as the reference category. Prior mathematics achievement was a consistent predictor of stronger profile membership. Each one-point increase in prior mathematics score increased the odds of membership in the developing procedural profile by 4% (OR = 1.04, 95% CI [1.01, 1.07], $p = 0.016$) and the odds of membership in the advanced balanced profile by 12% (OR = 1.12, 95% CI [1.09, 1.14], $p < 0.001$). Learners with a mathematics textbook at home also had higher odds of belonging to the developing procedural profile (OR = 1.55, 95% CI [1.01, 2.38], $p = 0.046$) and the advanced balanced profile (OR = 2.08, 95% CI [1.22, 3.57], $p = 0.008$). Similarly, learners drawn from private schools had higher odds of being in both the developing procedural profile (OR = 2.00, 95% CI [1.22, 3.27], $p = 0.006$) and the advanced balanced profile (OR = 1.63, 95% CI [1.18, 2.27], $p = 0.003$).

The predictors did not operate uniformly across the two stronger profiles. Data-task exposure predicted membership in the developing procedural profile (OR = 1.52, 95% CI [1.08, 2.15], $p = 0.017$), but not in the advanced balanced profile (OR = 0.97, 95% CI [0.53, 1.79], $p = 0.925$). This suggests that exposure to data-related mathematics tasks may help learners move beyond the weakest reasoning profile but may not be sufficient for advanced balanced reasoning unless such tasks require deeper interpretation, comparison, and evidence-based justification. Rural school location was associated with lower odds of belonging to the advanced balanced profile (OR = 0.44, 95% CI [0.22, 0.88], $p = 0.021$) but did not significantly predict membership in the developing procedural profile. Peri-urban location and sex were not significant predictors in either comparison. These findings indicate that profile membership was most strongly

 Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

associated with prior mathematics achievement, access to learning resources, school context, and the nature of learners' exposure to data-related tasks.

Table 6: Multinomial logistic regression predicting profile membership

Comparison	Predictor	OR	95% CI	*p*-value
Developing vs emerging	Prior mathematics score	1.04	[1.01, 1.07]	0.016
	Textbook at home (yes)	1.55	[1.01, 2.38]	0.046
	Private school	2.00	[1.22, 3.27]	0.006
	Peri-urban location	1.22	[0.72, 2.07]	0.453
	Rural location	0.82	[0.47, 1.43]	0.489
	Female	1.44	[0.86, 2.4]	0.167
	Data-task exposure	1.52	[1.08, 2.15]	0.017
Advanced vs emerging	Prior mathematics score	1.12	[1.09, 1.14]	< 0.001
	Textbook at home (yes)	2.08	[1.22, 3.57]	0.008
	Private school	1.63	[1.18, 2.27]	0.003
	Peri-urban location	0.76	[0.56, 1.04]	0.085
	Rural location	0.44	[0.22, 0.88]	0.021
	Female	1.28	[0.84, 1.96]	0.247
	Data-task exposure	0.97	[0.53, 1.79]	0.925

Source: *Author's analysis of field data, 2026*

Table 7 shows that the retained profiles differed significantly across all three distal outcomes. The largest difference was observed for mathematics confidence ($F = 189.57$, $p < 0.001$, $\eta^2 = 0.466$), which indicates that profile membership accounted for 46.6% of the variance in learners' mathematics confidence. Confidence increased steadily from the emerging statistical reasoning profile ($M = 2.78$, $SD = 0.35$) to the developing procedural reasoning profile ($M = 3.20$, $SD = 0.35$) and was highest among learners in the advanced balanced reasoning profile ($M = 3.68$, $SD = 0.29$). Prior mathematics scores also differed significantly across profiles ($F = 30.07$, $p < 0.001$, $\eta^2 = 0.122$), with the advanced balanced reasoning profile recording the highest mean prior mathematics score. Classroom engagement showed a smaller but still significant difference ($F = 7.06$, $p < 0.001$, $\eta^2 = 0.032$).

These findings show that the statistical reasoning profiles had educational meaning beyond the profile indicators themselves. Learners in the advanced balanced reasoning profile were not only stronger in data representation, computation, interpretation, comparison, and evidence-based conclusion drawing, but also reported greater mathematics confidence and had stronger prior mathematics achievement. The comparatively large effect for mathematics confidence suggests that learners' beliefs about their mathematical capability may be closely linked to how they engage with statistical reasoning tasks, although the cross-sectional design does not support causal interpretation. The smaller effect for classroom engagement suggests that general engagement in mathematics lessons may be less distinctive than confidence and prior achievement in separating the profiles.

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 3, 2026 (pp. 1-27)
www.abjournals.org

Table 7: Distal outcomes by latent profile

Outcome	Emerging M (SD)	Developing M (SD)	Advanced M (SD)	F-Stat	*p-value*	η^2
Mathematics confidence	2.78 (0.35)	3.20 (0.35)	3.68 (0.29)	189.57	< 0.001	0.466
Classroom engagement	3.17 (0.35)	3.24 (0.37)	3.35 (0.33)	7.06	< 0.001	0.032
Prior mathematics score	52.06 (9.75)	55.53 (9.72)	62.44 (10.07)	30.07	< 0.001	0.122

Note. η^2 indicates the proportion of outcome variance associated with profile membership in the omnibus comparison.

Source: *Author's analysis of field data, 2026*

DISCUSSION

The study set out to determine whether JHS learners' statistical reasoning could be understood through distinct profiles rather than through a single total score. The findings support that position. The retained three-profile solution showed that learners differed not only in overall level of performance but also in the configuration of reasoning across representation, computation, interpretation, comparison, and evidence-based conclusion drawing. This result strengthens the argument that school-level statistical reasoning is multidimensional and that diagnostic assessment should preserve domain-specific information before it is collapsed into an overall mark.

The emerging statistical reasoning profile represented learners who showed weak evidence across all five domains. This profile is important because it identifies learners who may require broad foundational support, not only additional practice with calculations. The pattern suggests limited command of the representational, numerical, interpretive, and evidential resources needed to work with data. In classroom terms, these learners would likely benefit from structured tasks that begin with familiar contexts, concrete data, clear displays, guided interpretation, and oral or written explanations that connect answers to evidence.

The developing procedural reasoning profile is the most theoretically revealing finding because it was the largest group and showed a clear imbalance: learners were relatively stronger in computation and data representation but weaker in interpretation, comparison, and evidence-based conclusion drawing. This profile speaks directly to a central concern in statistics education: learners may learn statistics as procedures without developing statistical reasoning as sense-making with data. The finding is consistent with frameworks that distinguish representation and reduction of data from the more demanding work of interpretation, comparison, and inference (Jones et al., 2000; Mooney, 2002; Garfield & Ben-Zvi, 2008). It also aligns with research showing that learners' explanations and representations can develop unevenly and that interpreting patterns, variability, and evidence remains difficult even when learners can produce displays or numerical answers (Balkaya & Kurt, 2023; Oslington et al., 2024).

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

The advanced balanced reasoning profile showed comparatively strong performance across all five domains. Its importance lies not only in being the highest-performing profile but also in being more balanced. These learners appeared more able to coordinate data displays, numerical summaries, contextual interpretation, comparison, and evidence-based conclusion drawing. In theoretical terms, this profile approximates the kind of integrated statistical reasoning described in the literature: reasoning that moves beyond isolated values to aggregate patterns, connects representation to meaning, and uses data as evidence for claims (Makar & Rubin, 2009; Watson & Callingham, 2003; Wild & Pfannkuch, 1999).

The profile structure therefore makes an international contribution beyond the Ghanaian case. It demonstrates how a person-centred lens can reveal a common instructional problem in school statistics: the transition from procedural data handling to evidence-based statistical reasoning. A variable-centred analysis could show that learners score lower on interpretation or evidence-based conclusion drawing, but it would not show how these weaknesses cluster within learners. The profile approach shows that many learners are not simply "low" or "average"; rather, they may be procedurally competent in some respects but underdeveloped in the inferential and evidential aspects of reasoning.

The predictors of profile membership should be interpreted as contextual associations rather than causal effects. Prior mathematics achievement was the most consistent predictor of stronger profile membership, which is unsurprising because statistical reasoning requires numerical fluency, representational knowledge, and familiarity with mathematical tasks. However, the domain patterns show that statistical reasoning cannot be reduced to general mathematics achievement. Learners may have enough mathematical skill to compute or complete displays but still struggle to interpret, compare, or justify conclusions. This distinction is central to the practical value of the study.

Textbook access and private school attendance were associated with stronger profile membership, while rural school location reduced the odds of membership in the advanced balanced profile. These findings point to opportunity-to-learn differences, but they should not be read as deterministic claims about school type or location. Textbook access may proxy opportunities for revision and repeated exposure to structured tasks. Private school attendance may proxy differences in class routines, instructional supervision, resource availability, or home support. A rural location may reflect constraints in material resources, teacher deployment, class organisation or access to enrichment opportunities. Future research should examine these mechanisms directly through classroom observation and teacher interviews.

The finding on data-task exposure is especially instructive. Exposure predicted membership in the developing procedural profile but not the advanced balanced profile. This suggests that simply increasing the number of data-handling tasks may help learners move beyond the weakest profile, but it may not be enough to produce balanced statistical reasoning. Routine tasks involving tables, graphs, and averages can strengthen representation and computation; advanced reasoning requires tasks that ask learners to compare groups, examine variation, critique claims, and justify conclusions using data. In plain classroom language, more tasks are useful, but better-designed tasks are the leverage point.

The distal outcome results add educational meaning to the profiles. Learners in the advanced balanced profile reported the highest mathematics confidence and had the strongest prior

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

mathematics achievement, while learners in the emerging profile reported the lowest levels. Mathematics confidence showed the largest profile-related difference. This does not establish causal direction, but it suggests that learners' beliefs about their mathematical capability are closely connected with how they engage with statistical reasoning tasks. The smaller difference in classroom engagement may mean that general participation in mathematics lessons is less discriminating than confidence, prior achievement, or more targeted measures of cognitive engagement with interpretation and justification.

The instructional implication is that profile labels should be used diagnostically, not hierarchically. Learners in the emerging profile need carefully scaffolded experiences across all domains. Learners in the developing procedural profile need tasks that deliberately link calculation and representation to explanation, comparison, variability, and evidence. Learners in the advanced balanced profile need enrichment through more open data investigations, informal inference, critique of claims, and communication of uncertainty. For curriculum and teacher development, the findings suggest that data handling should be taught as reasoning with evidence, not as a checklist of graphing and calculation procedures.

The study has limitations that should guide further work. The cross-sectional design identifies profiles and associations but cannot explain how learners move between profiles over time. The municipal sample provides context-sensitive evidence for Bolgatanga Municipality but should not be generalised uncritically to all JHS learners in Ghana or to other African contexts. The assessment was intentionally aligned to JHS-level statistical reasoning, but no single instrument can capture the full richness of learners' thinking. Future studies should use longitudinal designs, classroom observations, teacher interviews, and learner response analyses to examine how instructional practices support movement from procedural data handling to balanced statistical reasoning.

CONCLUSION

This study examined latent profiles of statistical reasoning among JHS learners in Bolgatanga Municipality, Ghana. The findings show that learners' reasoning with data was not adequately represented by a single total score or broad achievement category. Latent profile analysis identified three statistically and instructionally meaningful profiles: emerging statistical reasoning, developing procedural reasoning, and advanced balanced reasoning. The largest profile was not simply weak; it was procedural, with comparatively stronger computation and data representation but weaker interpretation, comparison and evidence-based conclusion drawing. This pattern shows why school statistics assessment must distinguish data handling from statistical reasoning with evidence.

Learner- and school-level factors were associated with profile membership, including prior mathematics achievement, textbook access, school ownership, data-task exposure, and school location. The findings do not support causal claims, but they suggest that opportunities to learn and practise meaningful data tasks may shape learners' reasoning profiles. The study contributes to statistics education research by offering a theoretically linked and empirically grounded profile account of statistical reasoning in an under-represented African basic education context. Its practical message is direct: improving statistical reasoning requires more than asking learners to draw graphs and calculate averages. Learners also need repeated

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

opportunities to interpret data in context, compare patterns, evaluate claims, and justify conclusions with evidence.

AI Assistance Disclosure

AI assistance was used for language refinement, organisation of sections, and coherence checking during manuscript preparation. AI assistance was not used to collect data, conduct statistical analysis, determine the final analytic model, or make independent scholarly decisions. All statistical procedures, interpretations, citations, and conclusions remain the responsibility of the author.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association.

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modelling: Three-step approaches using Mplus. Structural Equation Modelling: A Multidisciplinary Journal, 21(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modelling with continuous distal outcomes. Structural Equation Modelling: A Multidisciplinary Journal, 23(1), 20-31.

Balkaya, T., & Kurt, G. (2023). Middle school students' statistical reasoning about distribution in their statistical modeling processes. Statistics Education Research Journal, 22(1). https://doi.org/10.52041/serj.v22i1.312

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). Pre-K-12 guidelines for assessment and instruction in statistics education II: A framework for statistics and data science education (2nd ed.). American Statistical Association.

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organisational research: Determining appropriate sample size in survey research. Information Technology, Learning, and Performance Journal, 19(1), 43-50.

Ben-Zvi, D., & Garfield, J. (Eds.). (2004). The challenge of developing statistical literacy, reasoning, and thinking. Kluwer Academic Publishers.

Berlin, K. S., Williams, N. A., & Parra, G. R. (2014). An introduction to latent variable mixture modelling: Overview and cross-sectional latent class and latent profile analyses. Journal of Pediatric Psychology, 39(2), 174-187. https://doi.org/10.1093/jpepsy/jst084

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7-74. https://doi.org/10.1080/0969595980050102

Bolgatanga Municipal Assembly. (n.d.). About us. https://bolma.gov.gh/about-us/

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning, and thinking (pp. 375-396). Kluwer Academic Publishers.

Cochran, W. G. (1977). Sampling techniques (3rd ed.). Wiley.

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

Collins, L. M., & Lanza, S. T. (2010). Latent class and latent transition analysis: With applications in the social, behavioural, and health sciences. Wiley.
Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British Journal of Psychology, 105(3), 399-412. https://doi.org/10.1111/bjop.12046
Dvir, M., & Ben-Zvi, D. (2018). The role of model comparison in young learners' reasoning with statistical models and modeling. ZDM Mathematics Education, 50(7), 1183-1196. https://doi.org/10.1007/s11858-018-0987-4
Dvir, M., & Ben-Zvi, D. (2023). The double-edged sword of conjecturing. Mathematical Thinking and Learning, 25(2), 153-176. https://doi.org/10.1080/10986065.2021.1940427
Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual Review of Psychology, 53, 109-132. https://doi.org/10.1146/annurev.psych.53.100901.135153
Enders, C. K. (2010). Applied missing data analysis. Guilford Press.
Ferguson, S. L., Moore, E. W. G., & Hull, D. M. (2020). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. International Journal of Behavioural Development, 44(5), 458-468. https://doi.org/10.1177/0165025419881721
Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). Guidelines for assessment and instruction in statistics education: A pre-K-12 curriculum framework. American Statistical Association.
Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59-109. https://doi.org/10.3102/00346543074001059
Fu, Y., & Qi, C. (2026). Latent profile analysis of eighth-grade students' algebraic reasoning in China: What characterizes students' algebraic reasoning at different levels. The Journal of Educational Research. Advance online publication.
Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. International Statistical Review, 70(1), 1-25. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x
Garfield, J. (2002). The challenge of developing statistical reasoning. Journal of Statistics Education, 10(3). https://doi.org/10.1080/10691898.2002.11910676
Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. Springer.
Ghana Statistical Service. (2021). 2021 population and housing census: General report, volume 3A: Population of regions and districts. Ghana Statistical Service.
Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. Tutorials in Quantitative Methods for Psychology, 8(1), 23-34. https://doi.org/10.20982/tqmp.08.1.p023
Heritage, M. (2010). Formative assessment: Making it happen in the classroom. Corwin.
Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). Multilevel analysis: Techniques and applications (3rd ed.). Routledge.
Israel, G. D. (1992). Determining sample size. University of Florida Cooperative Extension Service.
Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. Mathematical Thinking and Learning, 2(4), 269-307. https://doi.org/10.1207/S15327833MTL0204_3

Article DOI: 10.52589/AJMSS-P3QSH4EW
DOI URL: https://doi.org/10.52589/AJMSS-P3QSH4EW

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kulesz, P. A., Roberts, G. J., Francis, D. J., Cirino, P. T., Walczak, M., & Vaughn, S. (2024). Latent profiles as predictors of response to instruction for students with reading difficulties. Journal of Educational Psychology, 116(3), 363-376.

Kurnia, A. B., Lowrie, T., & Patahuddin, S. M. (2024). The development of high school students' statistical literacy across grade levels. Mathematics Education Research Journal, 36, 7-35.

Laursen, B., & Hoff, E. (2006). Person-centred and variable-centred approaches to longitudinal data. Merrill-Palmer Quarterly, 52(3), 377-389.

Lehrer, R., & Schauble, L. (2004). Modelling natural variation through distribution. American Educational Research Journal, 41(3), 635-679. https://doi.org/10.3102/00028312041003635

Little, R. J. A., & Rubin, D. B. (2019). Statistical analysis with missing data (3rd ed.). Wiley.

Lohr, S. L. (2021). Sampling: Design and analysis (3rd ed.). CRC Press.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. Psychological Methods, 10(1), 21-39. https://doi.org/10.1037/1082-989X.10.1.21

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. Statistics Education Research Journal, 8(1), 82-105.

Masyn, K. E. (2013). Latent class analysis and finite mixture modelling. In T. D. Little (Ed.), The Oxford handbook of quantitative methods in psychology (Vol. 2, pp. 551-611). Oxford University Press.

McDonald, R. P. (1999). Test theory: A unified treatment. Lawrence Erlbaum Associates.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. Biochemia Medica, 22(3), 276-282.

Mooney, E. S. (2002). A framework for characterizing middle school students' statistical thinking. Mathematical Thinking and Learning, 4(1), 23-63. https://doi.org/10.1207/S15327833MTL0401_2

National Council for Curriculum and Assessment. (n.d.). Development of the Ghanaian curriculum. https://nacca.gov.gh/curriculum/

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modelling: A Monte Carlo simulation study. Structural Equation Modelling: A Multidisciplinary Journal, 14(4), 535-569. https://doi.org/10.1080/10705510701575396

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. Translational Issues in Psychological Science, 4(4), 440-461. https://doi.org/10.1037/tps0000176

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. Structural Equation Modelling: A Multidisciplinary Journal, 26(6), 967-985. https://doi.org/10.1080/10705511.2019.1590146

Oslington, G., Mulligan, J., & Van Bergen, P. (2024). Shifts in students' predictive reasoning from data tables in years 3 and 4. Mathematics Education Research Journal, 36, 547-576. https://doi.org/10.1007/s13394-023-00460-2

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. Journal of Educational Psychology, 86(2), 193-203. https://doi.org/10.1037/0022-0663.86.2.193

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). Knowing what students know: The science and design of educational assessment. National Academy Press.

Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: A classroom study. Mathematical Thinking and Learning, 13(1-2), 27-46. https://doi.org/10.1080/10986065.2011.538302

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Sage.

Revelle, W., & Condon, D. M. (2019). Reliability from alpha to omega: A tutorial. Psychological Assessment, 31(12), 1395-1411. https://doi.org/10.1037/pas0000754

Ruf, V., Thüs, D., Malone, S., Küchemann, S., Becker-Genschow, S., Vogel, M., Brünken, R., & Kuhn, J. (2026). A systematic review of empirical research on graphing numerical data in K-12 STEM education. International Journal of Science and Mathematics Education, 24, Article 28.

Schreiter, S., Friedrich, A., Fuhr, H., Malone, S., Brünken, R., Kuhn, J., & Vogel, M. (2024). Teaching for statistical and data literacy in K-12 STEM education: A systematic review on teacher variables, teacher education, and impacts on classroom practice. ZDM-Mathematics Education, 56, 31-45. https://doi.org/10.1007/s11858-023-01531-1

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. Contemporary Educational Psychology, 60, Article 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester Jr. (Ed.), Second handbook of research on mathematics teaching and learning (pp. 957-1009). Information Age Publishing.

Sijtsma, K. (2009). On the use, misuse, and very limited usefulness of Cronbach's alpha. Psychometrika, 74(1), 107-120. https://doi.org/10.1007/s11336-008-9101-0

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. Educational Psychologist, 50(1), 1-13. https://doi.org/10.1080/00461520.2014.1002924

Snijders, T. A. B., & Bosker, R. J. (2012). Multilevel analysis: An introduction to basic and advanced multilevel modelling (2nd ed.). Sage.

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and how-to guide of its application within vocational behaviour research. Journal of Vocational Behaviour, 120, Article 103445. https://doi.org/10.1016/j.jvb.2020.103445

Wang, C. K. J., Liu, W. C., Nie, Y., Chye, Y. L. S., Lim, B. S. C., Liem, G. A., Tay, E. G., Hong, Y.-Y., & Chiu, C.-Y. (2017). Latent profile analysis of students' motivation and outcomes in mathematics: An organismic integration theory perspective. Heliyon, 3(5), Article e00308. https://doi.org/10.1016/j.heliyon.2017.e00308

Wang, M.-T., & Degol, J. L. (2014). Staying engaged: Knowledge and research needs in student engagement. Child Development Perspectives, 8(3), 137-143. https://doi.org/10.1111/cdep.12073

Watson, J. M. (2006). Statistical literacy at school: Growth and goals. Lawrence Erlbaum Associates.

Watson, J. M., & Callingham, R. A. (2003). Statistical literacy: A complex hierarchical construct. Statistics Education Research Journal, 2(2), 3-46.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. Contemporary Educational Psychology, 25(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223-248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

World Bank. (2024). Confronting the learning crisis: Lessons from World Bank support for basic education, 2012-22. World Bank.

Xu, J., & Corno, L. (2022). A person-centred approach to understanding self-regulation in homework using latent profile analysis. Educational Psychology, 42(6), 767-786.

Zhang, L., Jiang, Y., & Lu, L. (2024). Adolescent students' psychological needs and educational outcomes in school: A latent profile analysis. The Journal of Experimental Education, 94(2), 284-302. https://doi.org/10.1080/00220973.2024.2378846