Arus Jurnal Pendidikan (AJP)
Website: http://jurnal.ardenjaya.com/index.php/ajpp
Email: jurnal.ajpp@ardenjaya.com

Enhancing Statistical Thinking through Cooperative Mathematics Learning on Data Dispersion: A Classroom Action Research Study (2025)

INFO PENULIS

Supratman
Universitas Sembilanbelas November Kolaka
supratmanmathusnkolaka@gmail.com

Ahmad Rustam
Universitas Sulawesi Tenggara
Ahmad.rustam1988@gmail.com

Andi Mariani Ramlan
Universitas Sembilanbelas November Kolaka
marianiramlan@gmail.com

La Ode Sirad
Universitas Sembilanbelas November Kolaka
laodesirad.usnkolaka@gmail.com

INFO ARTIKEL

ISSN: 2807-9558
Vol. 5, No. 3 Desember 2025
http://jurnal.ardenjaya.com/index.php/ajup

Saran Penulisan Referensi:

Supratman., Rustam, A., Ramlan, A. M., & Sirad, L. O. (2025). Enhancing Statistical Thinking through Cooperative Mathematics Learning on Data Dispersion: A Classroom Action Research Study (2025). *Arus Jurnal Pendidikan, 5*(3), 454-476.

Abstract

Statistical thinking is a fundamental competency in mathematics education, particularly in enabling students to interpret, analyze, and make decisions based on data. One essential topic that supports the development of statistical thinking is data dispersion, which includes measures such as range, variance, standard deviation, and interquartile range. However, many elementary students experience difficulties in understanding these concepts due to abstract explanations and teacher-centered instructional practices. This study aims to enhance students' statistical thinking skills through the implementation of cooperative learning in mathematics instruction focusing on data dispersion. The research employed Classroom Action Research (CAR) conducted in 2025 with fourth-grade students at a public elementary school in Indonesia. The study was implemented across three cycles, each consisting of planning, action, observation, and reflection stages. Data were collected through statistical thinking tests, observation sheets, and learning achievement assessments. The findings reveal a significant improvement in students' statistical thinking abilities and learning outcomes across cycles, with classical mastery increasing from 61.1% in Cycle I to 83.3% in Cycle III. The results indicate that cooperative mathematics learning effectively facilitates conceptual understanding of data dispersion and fosters active engagement, reasoning, and collaboration among students. This study contributes to mathematics education research by providing empirical evidence on the role of cooperative learning in strengthening statistical thinking at the elementary level.

Keywords: Statistical Thinking, Cooperative Learning, Mathematics Education, Data Dispersion, Classroom Action Research

A. Introduction

In the era of data-driven decision-making, statistical literacy and statistical thinking have become essential competencies for students at all levels of education. Statistical thinking enables learners to understand data variability, interpret patterns, and draw meaningful conclusions from data representations. The OECD emphasizes that statistical reasoning and data literacy constitute fundamental components of twenty-first-century skills that must be developed from early educational stages (OECD, 2019). As society becomes increasingly reliant on data for informed decision-making across various domains from public health to economics the ability to critically analyze and interpret statistical information has transformed from a specialized skill into a fundamental literacy requirement for all citizens (Gal, 2002; Watson, 2006). Statistical literacy extends beyond mere calculation of statistical measures; it encompasses the capacity to critically evaluate data sources, recognize patterns and trends, question statistical claims, and communicate data-based findings effectively (Rumsey, 2002). Ben-Zvi and Garfield (2004) argue that developing statistical thinking involves understanding why and how statistical investigations are conducted, recognizing the omnipresence of variability in data, and appreciating the role of statistical models in representing reality. Consequently, mathematics education must prioritize not only procedural knowledge but also conceptual understanding and reasoning within statistical contexts.

Within mathematics education, one of the foundational topics that supports statistical thinking is data dispersion. Measures of dispersion including range, variance, standard deviation, and interquartile range help students understand data variability beyond measures of central tendency such as mean, median, and mode (Shaughnessy, 2007). Understanding variability is central to statistical reasoning because it allows students to recognize that data distributions are not merely characterized by their centers but also by how spread out or clustered the values are (Reading & Shaughnessy, 2004). Research indicates that students often develop robust conceptions of central tendency relatively early, yet struggle significantly with variability concepts (Garfield & Ben-Zvi, 2008). This difficulty stems partly from the abstract nature of dispersion measures and the cognitive demands required to coordinate multiple data values simultaneously (delMas & Liu, 2005). Furthermore, students frequently exhibit misconceptions such as conflating variability with range exclusively, failing to recognize that variability exists even when datasets share the same range, or ignoring distributional shape when assessing spread (Makar & Confrey, 2005).

Despite its importance, research indicates that elementary students often struggle to grasp dispersion concepts due to their abstract nature and the dominance of procedural teaching approaches (Cobb, 1999; Shaughnessy, 2007). Traditional instruction that emphasizes formula application without conceptual grounding fails to develop students' intuitive understanding of why variability matters and how different measures capture distinct aspects of data spread (Garfield & Ben-Zvi, 2008). Consequently, students may mechanically calculate standard deviations without comprehending what these values reveal about data distribution or how they inform real-world interpretations. Observations conducted in a fourth-grade classroom revealed that mathematics instruction was predominantly teacher-centered, with limited opportunities for discussion, exploration, and collaborative reasoning. This pedagogical approach aligns with what Stigler and Hiebert (2009) describe as characteristic of many traditional mathematics classrooms, where teachers demonstrate procedures, students practice through repetition, and assessment focuses primarily on procedural accuracy rather than conceptual understanding. As a result, students demonstrated low engagement and superficial understanding of statistical concepts, particularly data dispersion.

The consequences of such instruction are well-documented in mathematics education research. Students taught through predominantly expository methods tend to develop fragile knowledge structures that do not transfer to novel contexts, exhibit lower motivation and mathematical self-efficacy, and demonstrate limited capacity for mathematical reasoning and problem-solving (Boaler, 2008; Hiebert & Grouws, 2007). In the context of statistical education specifically, passive learning environments inhibit the development of statistical thinking because students have insufficient opportunities to grapple with authentic data, engage in statistical argumentation, or develop informal inferential reasoning (Makar & Rubin, 2009). To address this issue, innovative instructional strategies that emphasize active learning and social interaction are required. Contemporary theories of mathematics learning, including constructivism and sociocultural perspectives, emphasize that meaningful mathematical understanding emerges through active engagement with mathematical tasks, social negotiation of meaning, and reflection on problem-solving processes (Vygotsky, 1978; von Glasersfeld, 1995). These theoretical frameworks suggest that instructional approaches facilitating collaborative knowledge construction may prove particularly effective for developing statistical thinking.

Cooperative learning is widely recognized as an effective pedagogical approach that promotes student engagement, conceptual understanding, and higher-order thinking (Johnson & Johnson, 2009; Slavin, 2015). Through structured group interactions, students are encouraged to explain ideas, negotiate meaning, and construct knowledge collaboratively. The theoretical foundations of cooperative learning draw from multiple sources, including Vygotsky's (1978) sociocultural theory, which posits that cognitive development occurs through social interaction within learners' zones of proximal development, and Piaget's (1985) constructivist theory, which emphasizes cognitive conflict and equilibration as mechanisms for conceptual change. Previous studies have shown that cooperative learning positively impacts mathematical understanding and reasoning across various mathematical domains (Davidson & Major, 2014; Zakaria, Chin, & Daud, 2010). Meta-analyses conducted by Capar and Tarim (2015) demonstrate that cooperative learning approaches yield significant improvements in mathematics achievement compared to traditional instruction, with effect sizes ranging from moderate to large across different age groups and mathematical topics. Furthermore, cooperative learning has been associated with enhanced problem-solving skills, increased mathematical communication, and improved attitudes toward mathematics (Hossain & Tarmizi, 2013).

In the context of statistical education, cooperative learning holds particular promise because statistical reasoning inherently involves discussion, argumentation, and collective sense-making of data (Chance, 2002). When students work collaboratively on statistical tasks, they must articulate their reasoning, justify their interpretations, consider alternative perspectives, and negotiate shared understanding processes that mirror authentic statistical practice (Wild & Pfannkuch, 1999). Specifically for understanding data dispersion, collaborative environments allow students to compare datasets, debate which distributional features matter most in different contexts, and develop more nuanced conceptions of variability through peer interaction (Bakker & Gravemeijer, 2004). However, empirical studies focusing specifically on statistical thinking and data dispersion at the elementary level remain limited. While research has examined cooperative learning effects on general mathematics achievement and on statistical concepts at secondary and tertiary levels (Garfield & Ben-Zvi, 2005; Roseth, Johnson, & Johnson, 2008), few studies have systematically investigated how cooperative instructional designs influence elementary students' development of variability concepts and statistical reasoning. This gap is particularly notable given that early experiences with data and variability establish foundational understandings that influence subsequent statistical learning (English, 2012).

The intersection of cooperative learning and elementary statistical education represents an underdeveloped area within mathematics education research. Specifically, there is limited empirical evidence regarding how cooperative learning structures can be effectively adapted to facilitate elementary students' understanding of data dispersion, what specific aspects of statistical thinking related to variability are most amenable to development through collaborative approaches, how teachers can design and implement cooperative learning activities that balance procedural skill development with conceptual understanding of dispersion measures, and what challenges emerge when implementing cooperative learning for statistical concepts and how these challenges can be addressed. Furthermore, much existing research on cooperative learning in mathematics has employed quantitative experimental designs that, while valuable for establishing general effectiveness, provide limited insight into the processes through which learning occurs and the contextual factors that mediate instructional outcomes (Gillies, 2016). Classroom action research offers a complementary methodological approach that enables teachers-as-researchers to systematically investigate instructional innovations within their own practice, generating situated knowledge about effective pedagogy while simultaneously improving student learning outcomes (Mertler, 2017).

Therefore, this study aims to investigate the effectiveness of cooperative mathematics learning in enhancing students' statistical thinking skills on the topic of data dispersion through a Classroom Action Research approach conducted in 2025. Specifically, the research seeks to examine how cooperative learning activities influence fourth-grade students' conceptual understanding of data dispersion measures, analyze changes in students' statistical reasoning abilities related to variability and data distribution, identify effective cooperative learning structures and instructional strategies for teaching data dispersion at the elementary level, and document challenges encountered during implementation and strategies developed to address them. The significance of this study is multifaceted. Theoretically, it contributes to the limited body of research examining cooperative learning applications in elementary statistical education, particularly regarding variability concepts that are foundational yet challenging for young learners. Practically, it provides educators with evidence-based insights into how collaborative instructional approaches can be designed and implemented to develop statistical thinking, offering concrete examples of activities, grouping strategies, and facilitation techniques. Methodologically, it demonstrates how classroom action research can serve as a powerful tool for teacher professional development and pedagogical innovation within mathematics education. By addressing the identified gaps through systematic investigation of cooperative learning for data dispersion at the elementary level, this study aims to advance both theoretical understanding of how students develop statistical thinking through social interaction and practical knowledge of effective instructional approaches for teaching variability concepts. Ultimately, the findings may inform curriculum development, teacher education, and instructional practice in ways that better prepare elementary students to become statistically literate citizens capable of critical engagement with data in an increasingly complex information landscape.

Statistical thinking is a higher-order cognitive ability that enables individuals to understand, analyze, and interpret data within meaningful contexts. Wild and Pfannkuch (1999) define statistical thinking as a way of reasoning that involves recognizing variability in data, understanding the processes of data collection and analysis, and making inferences based on statistical evidence. Unlike procedural statistical knowledge, which focuses primarily on performing calculations, statistical thinking emphasizes conceptual understanding of why and how statistical methods are applied (Chance, 2002). Ben-Zvi and Garfield (2004) identify three interrelated components of statistical thinking: understanding data variability and its influence on conclusions, integrating contextual knowledge with data analysis, and distinguishing between causal and correlational relationships. These components develop progressively through sustained engagement with diverse statistical situations. Moreover, statistical thinking requires a critical disposition, including questioning data, identifying patterns, and evaluating claims supported by statistical evidence with appropriate skepticism (Garfield & Ben-Zvi, 2008).

Closely related to statistical thinking is statistical literacy, a broader construct that encompasses the ability to understand, evaluate, and communicate statistical information encountered in everyday life (Gal, 2002). Watson (2006) characterizes statistical literacy as the capacity to interpret tables, graphs, and numerical information presented in the media, to comprehend fundamental concepts such as probability and variability, and to critically assess unsupported or misleading statistical claims. In an era dominated by data-driven information, statistical literacy has become an essential competency for informed citizenship and decision-making across domains such as health, economics, and public policy (Rumsey, 2002). The development of statistical thinking and literacy in elementary school students requires instructional approaches that prioritize active data exploration, meaningful discussion of statistical ideas, and strong connections to real-world contexts. Research indicates that students who engage in authentic statistical activities from an early age develop stronger conceptual foundations and more robust reasoning abilities than those whose learning is limited to computational procedures (English, 2012; Makar & Rubin, 2009).

A central concept underlying statistical thinking is data dispersion, which refers to the extent to which values in a dataset vary or spread around a central tendency. Shaughnessy (2007) emphasizes that variability is the core of statistical reasoning, as nearly all real-world phenomena exhibit variation, and ignoring this variation can lead to flawed interpretations. Common measures of dispersion, including range, variance, standard deviation, and interquartile range, provide different perspectives on how data are distributed. Reading and Shaughnessy (2004) describe a progression of students' understanding of variability, beginning with basic awareness that data vary, advancing through verbal descriptions, quantitative measurement, and ultimately the use of variability concepts for statistical inference. However, research shows that many students struggle to move beyond descriptive levels of understanding due to limited opportunities to engage meaningfully with data (delMas & Liu, 2005).

Empirical studies have documented persistent misconceptions related to dispersion concepts. Students often equate range with overall variability, assuming that datasets with identical ranges exhibit similar dispersion despite markedly different distributional patterns (Makar & Confrey, 2005). Additional difficulties include understanding that datasets with the same mean can have different levels of variability and an overemphasis on extreme values at the expense of considering the overall distribution (Garfield & Ben-Zvi, 2008). Standard deviation, in particular, is frequently treated as a formula to be memorized rather than as an intuitive measure describing how typical values deviate from the mean (Cooper & Shore, 2010). Addressing these misconceptions requires instructional strategies that emphasize visual representations, comparative analysis of distributions, and contextualized data exploration. The use of dot plots, box plots, and histograms, as well as tasks involving distributions with equal means but differing dispersions, has been shown to support deeper conceptual understanding of variability (Bakker & Gravemeijer, 2004; Cobb, 1999).

Constructivist learning theory provides a strong theoretical foundation for understanding how students develop statistical concepts. From a constructivist perspective, knowledge is actively constructed by learners through interaction with experiences rather than passively received from instruction (von Glasersfeld, 1995). In statistics education, this implies that students must actively engage with data, test conjectures, and revise their understanding based on evidence in order to meaningfully grasp concepts such as dispersion. Piaget (1985) highlights the importance of cognitive conflict or disequilibrium in learning, which arises when existing cognitive structures fail to explain new phenomena. For example, encountering datasets with identical ranges but different dispersion patterns can challenge students' initial conceptions and prompt conceptual reorganization. Teachers play a critical role in facilitating this process by designing tasks that are cognitively challenging yet accessible and by supporting reflection and sense-making (Driver & Bell, 1986).

Social constructivist perspectives further emphasize the role of interaction in learning. Vygotsky (1978) argues that learning occurs first at the social level through interaction with others and is later internalized at the individual level. The concept of the zone of proximal development suggests that students learn most effectively when they engage in tasks slightly beyond their independent capabilities with appropriate support. Cooperative learning environments provide opportunities for such scaffolding, enabling students to articulate their reasoning, negotiate meaning, and internalize more sophisticated statistical ideas. Cobb and Yackel (1996) integrate individual and social constructivist perspectives by emphasizing the importance of sociomathematical norms, through which classroom communities establish shared expectations about valid explanations, arguments, and solutions. This perspective is particularly relevant to statistical thinking, which involves reasoning, argumentation, and interpretation rather than mere computation.

Cooperative learning has been widely recognized as an effective instructional approach in mathematics and statistics education. Johnson and Johnson (2009) identify essential elements of cooperative learning, including positive interdependence, individual accountability, promotive interaction, social skills, and group processing. From a motivational perspective, Slavin (2015) argues that cooperative learning structures that combine group rewards with individual accountability encourage students to support one another's learning. Meta-analytical evidence indicates that cooperative learning has a significant positive effect on mathematics achievement across educational levels and contexts (Capar & Tarim, 2015). Beyond academic achievement, cooperative learning has been shown to improve students' attitudes toward mathematics, social relationships, and engagement in learning (Johnson et al., 2000; Roseth et al., 2008).

In mathematics classrooms, cooperative learning facilitates meaningful mathematical discourse, allowing students to articulate their reasoning, consider alternative perspectives, and refine their understanding through dialogue (Davidson & Major, 2014). The quality of interaction within cooperative groups, particularly the provision of elaborated explanations, has been found to be a strong predictor of learning gains (Webb, 1991). These benefits are especially salient in statistics education, where reasoning about data is inherently social and interpretative. Research suggests that statistical reasoning, particularly about variability, develops most effectively when students engage in group discussions, hands-on data investigations, and collective reflection on statistical meaning (Chance, 2002; Garfield & Ben-Zvi, 2005).

At the elementary level, the development of statistical thinking requires instructional approaches that align with students' cognitive development while maintaining conceptual rigor. Studies demonstrate that young learners are capable of engaging in sophisticated statistical reasoning when tasks are designed appropriately and supported by concrete and visual representations (English, 2012). The use of dynamic data visualization tools enables students to explore distributions and variability in intuitive ways, supporting discovery-based learning (Konold & Higgins, 2003). Approaches that involve students in the full cycle of statistical inquiry—from posing questions and collecting data to analyzing and communicating results—have been shown to foster deep understanding of variability and inference (Lehrer & Schauble, 2004). Importantly, research indicates that progression in students' understanding of averages, variability, and graphical representations depends more on the quality of instructional experiences than on age alone (Watson & Moritz, 2000; Friel et al., 2001). Consequently, instructional designs that integrate cooperative learning, constructivist principles, and meaningful engagement with data are well positioned to support the development of statistical thinking, particularly in relation to concepts of data dispersion and variability in elementary education.

B. Methodology

1. Research Design

This study employed Classroom Action Research (CAR) as its methodological framework to investigate the effectiveness of cooperative mathematics learning in enhancing students' statistical thinking skills on data dispersion. Classroom Action Research is a systematic approach that enables teachers to investigate and improve their own practice through iterative cycles of planning, acting, observing, and reflecting (Kemmis & McTaggart, 2005). This methodology was selected because it allows for continuous refinement of instructional strategies based on ongoing analysis of student learning and classroom dynamics, making it particularly appropriate for developing and testing innovative pedagogical approaches in authentic educational contexts.

The research was conducted through three complete action research cycles during the 2025 academic year, with each cycle building upon insights gained from the previous one. Each cycle consisted of four interconnected stages: planning, action, observation, and reflection. The planning stage involved designing instructional activities and materials based on identified needs and previous cycle outcomes. The action stage consisted of implementing the planned cooperative learning interventions in the classroom. The observation stage involved systematic data collection to document both teaching processes and student learning. The reflection stage included analyzing collected data to evaluate the effectiveness of interventions and identify areas for improvement in subsequent cycles. This cyclical structure allowed for progressive refinement of cooperative learning strategies and increasingly targeted support for students' development of statistical thinking about data dispersion.

The first cycle focused on introducing basic cooperative structures and fundamental concepts of data dispersion, establishing classroom norms for collaborative work, and assessing students' initial understanding. The second cycle built upon the foundation established in Cycle 1 by introducing more sophisticated cooperative learning activities and deepening conceptual understanding of variability measures. The third cycle emphasized consolidation of learning, application of statistical thinking to novel contexts, and development of students' ability to communicate statistical reasoning.

2. Research Setting and Participants

The research was conducted at SMP Negeri 1 Tanggetada, a public junior high school in kolaka, Indonesia, during the second semester of the 2024/2025 academic year. The school serves a diverse student population representing various socioeconomic backgrounds and academic ability levels, making it a representative setting for investigating instructional innovations intended for broader application in Indonesian secondary education contexts.

The participants consisted of 18 fourth-grade students (9 males and 9 females) aged 9 to 10 years old enrolled in one intact classroom. The selection of this particular classroom was based on purposive sampling, specifically chosen because preliminary observations and assessment data indicated that students demonstrated particular challenges with statistical concepts, particularly in understanding data variability and dispersion measures, and would benefit from instructional innovation. These challenges were evidenced by low performance on prior assessments involving data interpretation and superficial engagement during traditional teacher-centered mathematics instruction.

The student group was heterogeneous in terms of mathematical ability, with prior mathematics achievement ranging from below-average to above-average based on previous semester grades and teacher assessments. This heterogeneity was advantageous for the study because cooperative learning theory suggests that mixed-ability groups provide opportunities for peer scaffolding and benefit students across the achievement spectrum. Prior to the commencement of the study, informed consent was obtained from school administrators, parents/guardians of all participating students, and assent was obtained from students themselves.

To facilitate cooperative learning, students were organized into six heterogeneous groups of three members each. Group composition was carefully designed based on multiple criteria: academic ability, gender balance, social dynamics, and communication skills. Group assignments remained consistent across all three cycles to allow students to develop productive working relationships, though minor adjustments were made between cycles when observation data indicated that particular group configurations were not functioning effectively.

3. Instructional Intervention

The instructional intervention consisted of systematically designed cooperative mathematics learning activities focused on developing students' understanding of data dispersion and related statistical thinking skills. The intervention incorporated several evidence-based cooperative structures including Student Teams-Achievement Divisions (STAD), Think-Pair-Share, Jigsaw, and Numbered Heads Together. These structures were flexibly combined and adapted based on the specific learning objectives of each lesson and the developmental needs of fourth-grade students.

The instructional content was sequenced across the three action research cycles as follows. Cycle 1 introduced foundational concepts of data variability and range, including collecting and organizing data, creating visual representations, calculating range, and discussing what range reveals about data variability. Cycle 2 focused on understanding distribution shape and interquartile range, including analyzing datasets with different distribution patterns, dividing datasets into quartiles, calculating IQR, creating box plots, and comparing different dispersion measures. Cycle 3 introduced standard deviation concepts in developmentally appropriate ways and emphasized integrated statistical thinking through complete investigations involving data collection, analysis with multiple measures, and communication of findings.

Instructional materials included manipulatives for hands-on data collection, graph paper and templates for creating visual representations, data cards for physical manipulation, real datasets from contexts meaningful to students, cooperative learning role cards, thinking prompts to guide discussions, and technology tools when available. The teacher-researcher played facilitative roles including establishing cooperative norms, monitoring group interactions, providing scaffolding, orchestrating mathematical discourse, and conducting formative assessment.

4. Data Collection Instruments

Multiple data collection methods were employed to comprehensively investigate the effectiveness of cooperative learning. A Statistical Thinking Test was developed to assess students' conceptual understanding and reasoning abilities related to data dispersion. The test consisted of 15 items designed to assess conceptual understanding, procedural knowledge, interpretation skills, comparison and reasoning abilities, and application of concepts to real-world scenarios. The test was administered as a pre-test before Cycle 1 and as post-tests after each cycle. Test items were validated through expert review and pilot testing, with Cronbach's alpha of 0.82 indicating good internal consistency.

Observation sheets documented student and teacher activities during instructional sessions. The Student Activity Observation Sheet recorded participation in group discussions, collaborative behaviors, on-task behavior, communication patterns, problem-solving approaches, social interactions, and use of statistical language. The Teacher Activity Observation Sheet documented types of instructional activities, quality of explanations, questioning techniques, scaffolding strategies, feedback provided, and classroom management. Observations were recorded using systematic time-sampling approaches.

Learning Achievement Tests were administered at the end of each cycle to assess students' procedural proficiency and problem-solving abilities. These tests required students to calculate dispersion measures accurately, interpret data representations, select appropriate measures, and solve word problems. Mastery was defined as achieving a score of 75% or higher, consistent with Indonesian education standards.

Additional data sources included the teacher's reflection journal documenting observations, insights, and evolving understanding after each lesson; student work artifacts such as completed worksheets, graphs, written explanations, and presentations; and audio recordings of selected group discussions to capture the quality of peer interaction and statistical discourse.

5. Data Analysis Procedures

Data analysis was ongoing and cyclical, occurring both within each cycle formatively and across cycles summatively. Quantitative data from tests were analyzed using descriptive statistics including mean scores, standard deviations, score ranges, and distribution characteristics. Learning gains were calculated by comparing pre-test to post-test performance, with effect sizes computed to quantify the magnitude of improvement. The percentage of students reaching the mastery criterion of 75% was calculated for each cycle. Item analysis identified concepts showing greatest improvement and persistent difficulties.

Qualitative data from observations, reflection journals, transcripts, and student work were analyzed using thematic analysis procedures. Initial coding identified interesting features related to research questions and emergent themes. Codes were collated into potential themes representing patterns across the dataset. Themes were reviewed, defined, and interpreted in relation to research questions and theoretical frameworks.

Integration of quantitative and qualitative data occurred through triangulation to confirm interpretations, complementarity to explain quantitative patterns with qualitative insights, and development using findings from one data type to inform collection of another. At the conclusion of each cycle, a systematic reflection process integrated all data sources to evaluate effectiveness and plan improvements for the next cycle.

a. Results and Discussion

1. Results

This chapter presents the findings from three cycles of classroom action research investigating the effectiveness of cooperative mathematics learning in enhancing students' statistical thinking skills on data dispersion. Data were collected through statistical thinking tests, learning achievement tests, classroom observations, teacher reflection journals, and student work artifacts. The findings are organized by research cycle, followed by cross-cycle analysis of patterns and trends.

a. Baseline Data (Pre-Test Results)

Prior to implementing the cooperative learning intervention, a pre-test was administered to assess students' initial understanding of data dispersion concepts and statistical thinking abilities. The pre-test results provided baseline data against which subsequent improvements could be measured and revealed specific areas of difficulty that informed instructional planning for Cycle 1.

The pre-test mean score was 58.3 (SD = 12.7) out of a possible 100 points, indicating that students' initial understanding of data dispersion was limited. Only 5 out of 18 students (27.8%) achieved the mastery criterion of 75% or above, while the remaining 13 students (72.2%) scored below this threshold. The distribution of scores ranged from a minimum of 40 to a maximum of 80, revealing considerable variability in students' prior knowledge and readiness for learning about data dispersion.

Item analysis of the pre-test revealed specific patterns in students' understanding and misconceptions. Students performed relatively well on basic items requiring identification of maximum and minimum values in datasets (mean item score: 78%) and simple calculation of range (mean item score: 65%). However, performance declined substantially on items requiring conceptual understanding of what dispersion measures reveal about data (mean item score: 42%), interpretation of dispersion in context (mean item score: 38%), and comparison of distributions with different variability characteristics (mean item score: 35%). Students demonstrated particular difficulty with items involving interquartile range and standard deviation, concepts they had not previously encountered in their mathematics instruction.

Qualitative analysis of pre-test responses revealed several common misconceptions and limited reasoning patterns. Many students conflated measures of center (mean, median) with measures of spread, stating, for example, that "the average tells us how spread out the data are." When asked to explain what it means for data to be "spread out" or "variable," most students provided vague or circular definitions such as "it means the numbers are different" without articulating how much or in what pattern the values differ. Few students spontaneously used visual representations to support their reasoning about data distributions, and those who did typically created simple lists or unsorted tables rather than graphs that reveal distributional shape.

These baseline findings confirmed the need for instructional intervention and provided specific targets for the cooperative learning activities implemented in subsequent cycles. The prevalence of procedural knowledge without conceptual understanding suggested that previous instruction had emphasized computation over reasoning, justifying the study's focus on collaborative sense-making activities that promote deeper engagement with statistical concepts.

b. Cycle I: Introduction to Data Variability and Range

Planning Phase

Based on pre-test results and preliminary observations, Cycle I was designed to introduce fundamental concepts of data variability and establish effective cooperative learning structures. The specific learning objectives for Cycle I were: (1) students can explain what data variability means in everyday language; (2) students can calculate range as a measure of spread; (3) students can interpret range in the context of real-world datasets; (4) students can compare datasets with different ranges and explain what differences reveal; and (5) students can work collaboratively in groups, demonstrating basic cooperative skills such as active listening, equitable participation, and respectful disagreement.

Instructional activities for Cycle I included four 90-minute lessons implemented over two weeks. Cooperative structures emphasized in this cycle were Think-Pair-Share for initial exploration of variability concepts and STAD (Student Teams-Achievement Divisions) for practice and consolidation. Lesson 1 focused on collecting classroom data (student heights, arm spans, number of siblings) and creating visual representations to observe variation. Lesson 2 introduced range through comparison of datasets with obviously different spreads. Lesson 3 provided practice calculating range from various representations (tables, dot plots, line plots) in collaborative groups. Lesson 4 engaged students in comparing and interpreting ranges across different contexts, culminating in group presentations of findings.

Action and Observation Phase

Implementation of Cycle I revealed both successes and challenges in establishing cooperative learning for statistical thinking. Observation data indicated that initial group work was somewhat chaotic, with students unsure of how to share tasks and responsibilities equitably. Average on-task behavior during the first lesson was 68%, with considerable variation across groups (range: 52-85%). However, on-task behavior improved steadily across the four lessons, reaching 82% by Lesson 4 as students became more familiar with cooperative structures and clearer about group roles.

Student participation in group discussions showed interesting patterns. Initially, higher-achieving students dominated conversations in most groups, with lower-achieving students primarily listening or engaging in off-task behavior. By Lesson 3, after explicit instruction on cooperative skills and assignment of rotating roles (facilitator, recorder, reporter), participation became more equitable. Tally marks recording verbal contributions showed that by Lesson 4, the ratio of contributions between highest and lowest contributors within groups had decreased from approximately 5:1 to 2:1, indicating more balanced participation.

The quality of statistical discourse in groups evolved noticeably across Cycle I. Early discussions were brief and procedural, focused primarily on "what's the biggest number?" and "what's the smallest number?" with minimal conceptual elaboration. By Lessons 3 and 4, groups increasingly engaged in more substantive conversations about what range reveals. For example, one group discussing test score ranges stated: "Class A has a bigger range, so some students did really good and some did really bad. Class B's range is smaller, so everybody did about the same." This comment, while informal, demonstrated emerging understanding that range captures something about consistency vs. variability in performance.

Teacher observation notes documented several effective instructional moves that promoted statistical thinking in cooperative groups. Asking groups to predict which dataset would have a larger range before calculating, then explain discrepancies between predictions and results, generated productive cognitive conflict. Providing sentence frames such as "The range tells us about the data" and "A larger range means " scaffolded students' articulation of conceptual understanding. Requiring groups to create visual representations before calculating range helped students develop intuitive understanding of spread.

Challenges observed during Cycle I included students' tendency to focus exclusively on extreme values while ignoring the overall distribution pattern, difficulty articulating why range might be misleading when outliers are present, and some groups' struggles with equitable participation despite instructional support. Teacher reflection journal entries noted: "Students are beginning to understand variability conceptually but still default to procedural thinking when uncertain. Need to continue emphasizing 'what does this tell us?' rather than just 'how do we calculate this?'"

Reflection and Assessment Phase

The Cycle I post-test, administered after the fourth lesson, revealed substantial improvement from baseline. The mean score increased to 68.9 (SD = 11.4), representing a gain of 10.6 points from the pre-test (effect size d = 0.85, indicating a large effect). Eleven out of 18 students (61.1%) achieved the mastery criterion of 75% or above, compared to only 27.8% on the pre-test. This represented a 33.3 percentage point increase in classical mastery, though the target of 75% class-wide mastery was not yet achieved.

Item analysis showed that improvement was most pronounced on items requiring calculation of range (mean item score improved from 65% to 88%) and basic interpretation of range in context (improved from 38% to 64%). Students continued to struggle with more complex items requiring comparison of multiple distributions (improved from 35% to 48%) and explaining limitations of range as a measure (new item introduced in Cycle I post-test: 41% correct).

Analysis of student work artifacts from Cycle I revealed growing sophistication in visual representation of data. By Lesson 4, most groups spontaneously created ordered dot plots or line plots that clearly showed distribution patterns, whereas initial representations tended to be unordered lists. Written explanations of range interpretations remained relatively brief but showed increasing use of appropriate vocabulary such as "spread," "variability," "minimum," "maximum," and "consistent."

Observation data summarized across Cycle I indicated positive trends in cooperative behaviors. The percentage of students demonstrating active listening (indicated by eye contact, nodding, asking clarifying questions) increased from 56% in Lesson 1 to 78% in Lesson 4. The percentage of groups where all members contributed at least one substantive comment during observed discussion segments increased from 33% (2 out of 6 groups) to 67% (4 out of 6 groups). Instances of students providing explanations to peers (rather than just answers) increased from an average of 2.3 per lesson to 5.7 per lesson.

Reflection on Cycle I outcomes identified several implications for Cycle II. Strengths to maintain included the use of real, student-generated data that increased engagement and relevance; Think-Pair-Share structure that ensured individual thinking time before group discussion; and explicit instruction on cooperative skills integrated with mathematical content.

Areas requiring adjustment included the need for more scaffolding to help students articulate limitations of range and prepare them for more sophisticated dispersion measures; more structured protocols to ensure equitable participation in groups still struggling with dominance patterns; and increased emphasis on connections between visual representations and numerical measures.

c. Cycle II: Understanding Distribution Shape and Interquartile Range

Planning Phase

Building on Cycle I foundations, Cycle II aimed to deepen students' understanding of data dispersion by introducing concepts of distribution shape, quartiles, and interquartile range (IQR). The cycle was designed to address the limitation that range, while intuitive, does not capture information about how data are distributed between extremes or account for the impact of outliers. Specific learning objectives for Cycle II were: (1) students can explain that datasets with the same range may have different distribution patterns; (2) students can divide datasets into quartiles and identify the median; (3) students can calculate interquartile range and explain what it represents; (4) students can create and interpret box plots; (5) students can compare range and IQR and articulate situations where each measure is more informative; and (6) students demonstrate improved collaborative skills including constructive disagreement and synthesis of multiple perspectives.

Instructional activities for Cycle II consisted of five 90-minute lessons implemented over 2.5 weeks. The extended duration reflected the greater complexity of content and the need for substantial practice with new concepts. Cooperative structures emphasized in this cycle included Jigsaw for initial exploration of quartiles (with each student becoming an "expert" on one quartile), STAD for calculating IQR and creating box plots, and Numbered Heads Together for checking understanding and ensuring individual accountability.

Lesson 1 engaged students in comparing datasets with identical ranges but very different distribution patterns (e.g., uniform vs. clustered vs. bimodal), prompting discussion about what range fails to capture. Lesson 2 introduced the concept of dividing data into four equal parts and identifying quartile boundaries. Lesson 3 focused on calculating IQR as Q3 - Q1 and interpreting what IQR reveals about the middle 50% of data. Lesson 4 introduced box plots as visual representations of five-number summaries and provided practice creating box plots from datasets. Lesson 5 engaged students in comparative analysis tasks requiring them to use both range and IQR to describe and compare distributions, developing decision-making skills about which measure is most appropriate for particular purposes.

Action and Observation Phase

Implementation of Cycle II revealed that students' increased familiarity with cooperative structures allowed instructional focus to shift more fully to mathematical content, though challenges with new conceptual material required continued attention to collaborative support. On-task behavior remained high throughout Cycle II, averaging 84% across all five lessons (range: 78-89% across lessons). This consistency suggested that cooperative learning structures had been successfully internalized and that students understood expectations for productive group work.

The Jigsaw structure employed in Lesson 2 proved particularly effective for introducing quartile concepts. After initial instruction, each group member worked with "expert groups" comprising students from other teams who were studying the same quartile. These expert groups collaborated to develop deep understanding of their assigned quartile position and how to identify it in various datasets. Students then returned to their home groups to teach their expertise to teammates. Observation notes documented high engagement during this activity, with students taking seriously their responsibility to understand material well enough to teach others. One student commented: "I had to really get it because my group was counting on me to explain Q2 [the median]." This accountability structure motivated thorough understanding.

However, observation data also revealed that the conceptual demands of quartiles and IQR generated considerable confusion initially. In Lesson 2, approximately 72% of groups required teacher intervention to clarify how to divide datasets into four equal parts when the number of data points was not evenly divisible by four. Some groups persisted in misconceptions even after initial explanation, requiring multiple rounds of scaffolding. Teacher questioning strategies proved crucial: asking students "How many data points should be in each quartile?" and "Does that match what you found?" prompted groups to identify and correct their own errors more effectively than direct correction.

By Lesson 3, when calculating IQR, students demonstrated solid procedural understanding but initially struggled with conceptual interpretation. When asked "What does IQR tell us about the data?", typical initial responses were vague: "It tells us about the middle part" or "It's the range of the middle." Through structured group discussions using sentence frames ("The IQR of means that the middle 50% of data values are spread over units"), understanding deepened. By the end of Lesson 3, most groups could articulate that smaller IQR indicates more consistency in the middle of the distribution while larger IQR indicates more variability among typical values.

Box plot creation in Lesson 4 generated enthusiastic engagement, with students appreciating the visual efficiency of representing five numbers (minimum, Q1, median, Q3, maximum) in one compact graph. However, several groups initially constructed box plots incorrectly, treating the median as the center of the box rather than recognizing it could be off-center when data were skewed. Peer explanation proved effective in addressing this misconception. In one observed exchange, a student explained to a confused teammate: "The median doesn't have to be in the middle of the box! It just shows where the actual middle data point is. If most numbers are lower, the median will be closer to Q1." This explanation, more accessible than formal instruction about skewness, helped several students grasp an important conceptual point.

Lesson 5's comparative analysis activities revealed growing sophistication in statistical reasoning. Groups discussed scenarios such as: "Two companies have the same average salary. Company A has a range of $80,000 and IQR of $15,000. Company B has a range of $45,000 and IQR of $30,000. What might this tell you about salary distributions in each company?" Most groups recognized that Company A likely had some very high or very low outlier salaries (explaining the large range) but typical employees' salaries were quite similar (explaining the small IQR), while Company B had no extreme outliers but more variability among typical salaries. This type of nuanced reasoning about multiple measures simultaneously demonstrated substantial growth in statistical thinking.

Observation of collaborative behaviors in Cycle II showed continued improvement in cooperative skills. The percentage of groups demonstrating constructive disagreement—situations where students respectfully challenged each other's thinking and worked together to resolve differences increased from isolated instances in Cycle I to regular occurrences in Cycle II (observed in 83% of groups during at least one Cycle II lesson). For example, one group debated whether two specific data points should be included in Q2 or Q3, with members presenting arguments, checking their reasoning against the definition of quartiles, and ultimately reaching consensus. Such interactions exemplified the type of collaborative knowledge construction that cooperative learning aims to foster.

Audio recordings of group discussions revealed increasingly sophisticated use of statistical vocabulary and more extended chains of reasoning. In Cycle I, most group discussions consisted of brief exchanges focused on procedural steps. In Cycle II, discussions often included multiple turns where students built on each other's ideas, asked probing questions, and developed collective explanations. For instance, one recorded discussion showed a group working through why IQR might be preferred to range when outliers are present, with students sequentially contributing ideas that ultimately led to a coherent group understanding that no single member had articulated initially.

Reflection and Assessment Phase

The Cycle II post-test demonstrated continued improvement in statistical thinking and achievement. The mean score increased to 76.4 (SD = 10.2), representing a gain of 7.5 points from Cycle I post-test and 18.1 points from the original pre-test (effect size from pre-test d = 1.52, a very large effect). Thirteen out of 18 students (72.2%) achieved the mastery criterion of 75% or above, an increase of 11.1 percentage points from Cycle I and 44.4 percentage points from pre-test. While this represented substantial progress, the target of 75% classical mastery remained slightly out of reach.

Item-level analysis revealed differential patterns of improvement across content areas. Students showed strong performance on procedural items requiring calculation of quartiles (mean item score: 81%) and IQR (mean item score: 78%), and on basic interpretation items (mean item score: 74%). Performance remained weaker on items requiring complex comparison of distributions using multiple measures simultaneously (mean item score: 63%) and on explanation items asking students to justify why one measure might be preferred over another in specific contexts (mean item score: 58%). These patterns suggested that while students had developed solid procedural skills and basic conceptual understanding, the highest levels of statistical reasoning—involving metacognitive awareness of when and why to use particular tools—required additional development.

Analysis of box plots created by students during Cycle II showed substantial growth in representational competence. In early lessons, 44% of student-created box plots contained at least one significant error (incorrect quartile positions, missing elements, or scaling issues). By Lesson 5, this error rate had decreased to 17%, with most errors being minor (such as imprecise alignment) rather than conceptually based. Students' written interpretations of box plots similarly improved, progressing from simple identification of five-number summary values to richer descriptions incorporating distribution shape and comparative judgments.

Observation data summarized across Cycle II indicated that cooperative learning structures were functioning effectively. Average on-task behavior (84%) exceeded the target of 75% established in planning. The percentage of students actively participating in group discussions (defined as contributing at least three substantive comments during observed 15-minute discussion segments) was 89%, compared to 67% in Cycle I. All six groups demonstrated at least some instances of higher-order collaborative behaviors such as asking each other why questions, requesting explanations rather than just answers, and building on each other's ideas to develop understanding that extended beyond any individual's initial contribution.

Teacher reflection journal entries from Cycle II noted significant progress but also highlighted continuing challenges: "Students have really embraced group work and are genuinely helping each other learn. The Jigsaw approach worked beautifully for quartiles—students took ownership of teaching their peers. However, I notice that the conceptual leap from 'here's how to calculate IQR' to 'here's when IQR is more useful than range' is still difficult for many. Cycle III needs to emphasize application and decision-making even more, helping students develop judgment about which statistical tools to use in various situations."

Reflection on Cycle II identified adjustments for Cycle III. Successful elements to continue included the Jigsaw structure for complex new content, the use of real-world scenarios that required interpretation and decision-making rather than just calculation, and the continued emphasis on visual representations alongside numerical measures. Areas for enhancement included providing more open-ended investigation tasks where groups must decide which measures to calculate rather than being told, increasing metacognitive reflection activities where students articulate their problem-solving strategies and decision-making processes, and introducing simplified concepts of standard deviation to round out students' repertoire of dispersion measures.

d. Cycle III: Standard Deviation and Integrated Statistical Thinking

Planning Phase

Cycle III was designed as a culminating experience that introduced standard deviation concepts (in developmentally appropriate ways) and emphasized integrated application of statistical thinking across all learned dispersion measures. The primary goals shifted from introducing new procedures to fostering flexible, strategic use of statistical tools and deep conceptual understanding of variability. Specific learning objectives for Cycle III were: (1) students can explain that standard deviation measures typical distance from the mean; (2) students can interpret standard deviation values in context; (3) students can select appropriate dispersion measures for different purposes and justify their choices; (4) students can conduct complete statistical investigations including question formulation, data collection, analysis with multiple measures, and communication of findings; (5) students demonstrate advanced collaborative skills including effective division of complex tasks and synthesis of group members' contributions into coherent final products.

Instructional activities for Cycle III consisted of six 90-minute lessons implemented over three weeks. The extended timeline allowed for a substantial culminating project. Cooperative structures emphasized in this cycle included STAD for introducing standard deviation concepts, Think-Pair-Share for metacognitive reflection on strategy selection, and extended collaborative project work for statistical investigations.

Lessons 1-2 introduced standard deviation through hands-on activities where students physically measured distances of data points from the mean using number lines, developing intuitive understanding that standard deviation captures "typical distance from average." Rather than teaching the complex formula, students used simplified computational approaches appropriate for small datasets or used technology (spreadsheet software when available) to calculate standard deviation, with instructional focus on interpretation rather than computation. Lessons 3-4 engaged students in comparative activities where they analyzed multiple datasets using the full range of learned measures (range, IQR, standard deviation), discussing which measures revealed similar or different information and which were most useful for particular interpretive purposes. Lessons 5-6 consisted of group statistical investigation projects where teams formulated research questions, collected data, analyzed data using appropriate measures and visualizations, and prepared presentations of findings for classmates.

Action and Observation Phase

Implementation of Cycle III revealed the culmination of students' development in both statistical thinking and cooperative learning skills. The hands-on approach to introducing standard deviation in Lessons 1-2 proved highly effective. Students used number lines where data points were represented as marks and the mean was indicated with a different color. They physically measured the distance of each data point from the mean, then calculated the average distance (with some simplification, using mean absolute deviation as an accessible approximation of standard deviation). This concrete experience made the abstract concept tangible. One student remarked: "Oh, so standard deviation is like asking 'how far away are the numbers usually?' That makes sense!" This intuitive understanding, while informal, provided a solid foundation for interpreting standard deviation values.

Observation notes documented high engagement during standard deviation activities, with students animated in their discussions about what constitutes "typical" distance. Groups debated questions such as "Should we include the really big distance when we calculate typical distance?" This debate naturally led to discussions about how standard deviation (unlike range) is not overly influenced by single outliers because it averages distances across all data points—a sophisticated conceptual insight that emerged from student discourse rather than direct instruction.

However, some students struggled with the transition from measuring physical distances on number lines to interpreting numerical standard deviation values calculated from formulas or technology. For several students, the concrete-to-abstract bridge remained incomplete, and they could work successfully with the physical model but became uncertain when presented with a calculated standard deviation like "SD = 8.3." Additional scaffolding through analogies proved helpful: "If the standard deviation is 8.3 points, that means most test scores are about 8-9 points away from the average some a bit more, some a bit less." This concrete interpretation helped students make sense of abstract numerical values.

Lessons 3-4's comparative analysis activities generated rich statistical discourse. Groups were presented with datasets described using multiple measures and asked to draw inferences about what the data patterns suggested. For example, one scenario provided: "Dataset A: Mean = 75, Range = 45, IQR = 12, SD = 7. Dataset B: Mean = 75, Range = 40, IQR = 25, SD = 13." Groups discussed what these measures revealed about similarities and differences in the distributions. Observation of these discussions showed that most groups recognized that both distributions had the same center (mean = 75) but different spreads, that Dataset A likely had more extreme outliers (larger range but smaller IQR and SD), while Dataset B had no extreme outliers but more variability among typical values (larger IQR and SD despite similar range).

These discussions demonstrated integration of knowledge across all three cycles—students were simultaneously considering multiple measures, relating numerical values to distribution characteristics, and using appropriate statistical language to articulate their reasoning. The quality of argumentation observed in Cycle III discussions far exceeded what was typical in Cycle I. Students regularly supported claims with specific reference to measures ("We know there are outliers because the range is big but the IQR and SD are small"), considered alternative interpretations ("Or maybe..."), and synthesized multiple perspectives to reach group consensus.

The culminating statistical investigation projects in Lessons 5-6 provided authentic contexts for applying learned concepts. Groups formulated diverse research questions relevant to their interests: comparing height variability between males and females in their class, analyzing consistency of different students' quiz scores across the semester, investigating variability in daily temperatures across seasons, and examining spread in time spent on homework across different subjects. Each group collected or gathered existing data, created multiple visual representations (dot plots, box plots), calculated multiple dispersion measures, and prepared presentations interpreting their findings.

Observation of project work revealed sophisticated collaborative practices. Groups naturally divided tasks based on members' strengths while ensuring everyone understood all aspects of the work. Planning discussions showed metacognitive awareness: "You're really good at making neat graphs, so maybe you do the box plots. I'll calculate the IQR and standard deviation. Then we can all work together on explaining what it means." Regular check-ins within groups ensured shared understanding: "Before we go further, let's make sure we all agree about what the data are showing." When disagreements arose, groups employed constructive resolution strategies, often returning to the data or their calculations to adjudicate between competing interpretations.

Project presentations in Lesson 6 demonstrated the learning outcomes of the entire intervention. Groups presented coherently, using appropriate statistical vocabulary, clearly explaining their research questions, methods, and findings. Most presentations included insightful interpretations, such as: "We found that male and female heights have almost the same range, but the IQR is bigger for males. This means that while both groups have some tall and short outliers, typical male heights are more spread out while females are more clustered near the average." Audience members (other student groups) asked thoughtful questions, demonstrating engagement with statistical content: "Did you consider whether outliers were affecting your results?" These interactions exemplified the statistical literacy that the intervention aimed to develop.

Teacher observation of presentations noted that students spontaneously used multiple representation types to support arguments, integrated numerical measures with visual displays, and demonstrated comfort discussing uncertainty and limitations of their analyses. Several groups acknowledged sample size limitations or recognized that their findings might not generalize beyond their classroom context showing emerging appreciation for statistical inference concepts that extend beyond the formal scope of elementary dispersion instruction.

Reflection and Assessment Phase

The Cycle III post-test, serving as the final assessment of the intervention's impact, demonstrated substantial learning gains. The mean score increased to 82.7 (SD = 9.1), representing a gain of 6.3 points from Cycle II, 13.8 points from Cycle I, and 24.4 points from the original pre-test (effect size from pre-test d = 2.21, an exceptionally large effect). Fifteen out of 18 students (83.3%) achieved the mastery criterion of 75% or above, surpassing the target of 75% classical mastery established in research planning. This represented a 22.2 percentage point increase from Cycle II and a 55.5 percentage point increase from pre-test.

Notably, the three students who did not achieve mastery at the 75% level all scored between 70-74%, indicating they were approaching mastery and had still made substantial progress from their pre-test scores (gains ranging from 18-22 points). None of the 18 students failed to show meaningful improvement across the intervention, though the magnitude of gains varied.

Item analysis of the Cycle III post-test revealed strong performance across all content areas. Students achieved high success rates on procedural items involving calculation of range (94%), IQR (89%), and interpretation of standard deviation (81%). Performance on complex conceptual items also showed marked improvement. Items requiring comparison of multiple distributions using multiple measures achieved 78% mean success rate, up from 63% in Cycle II and 48% in Cycle I. Items requiring explanation of when and why to use particular measures achieved 74% success rate, up from 58% in Cycle II. Items requiring interpretation of standard deviation in context (new in Cycle III) achieved 73% success rate, indicating that despite the conceptual complexity, most students developed functional understanding.

The highest-level items on the Cycle III post-test—which required students to conduct mini-analyses of novel datasets, select appropriate measures independently, and provide written explanations integrating multiple measures—showed 71% mean success rate. While this was the lowest category on the Cycle III test, it still represented strong performance on very challenging tasks that require synthesis and application of knowledge. These items effectively differentiated between students who had developed deep, flexible understanding versus those whose understanding remained more procedural or context-specific.

Analysis of student work artifacts from Cycle III projects revealed sophisticated statistical thinking. All six group projects included multiple appropriate visual representations, accurate calculations of dispersion measures, and interpretations that went beyond simple description to draw inferences about what patterns suggested. Five out of six projects explicitly discussed relationships between different measures (e.g., why range and IQR told different stories about their data), demonstrating integrated understanding. Four out of six projects acknowledged limitations or uncertainties in their analyses, showing emerging critical thinking about statistical conclusions.

Observation data summarized across Cycle III indicated that cooperative learning was functioning at high levels. On-task behavior averaged 88% (range: 84-92% across lessons), the highest of any cycle. All 18 students (100%) actively participated in group discussions during observed segments, with contribution patterns largely equitable within groups (the ratio between highest and lowest contributors averaged 1.8:1, indicating relatively balanced participation). Instances of higher-order collaborative behaviors such as asking why/how questions, building on each other's ideas, constructively challenging assumptions, and engaging in extended chains of reasoning occurred in 100% of groups during Cycle III observations, compared to 83% in Cycle II and 67% in Cycle I.

Analysis of audio recordings from Cycle III group discussions showed that statistical conversations had become natural and spontaneous rather than prompted by teacher questions. Students regularly used sophisticated statistical vocabulary without awkwardness: variability, distribution, consistency, typical distance, quartiles, outliers, skewed. Explanations given by students to peers often mirrored the structure of effective statistical reasoning: identifying the question, referencing relevant data/measures, drawing inferences, acknowledging limitations. In one particularly strong example, a student explained to teammates: "We can't just use the range here because there's that one really high score that's an outlier. The IQR is better because it shows us how spread out the middle students are without being affected by that one weird score."

Teacher reflection journal entries from Cycle III expressed satisfaction with students' growth: "The transformation in these students' understanding and engagement has been remarkable. They've gone from barely able to explain what 'spread' means to conducting independent investigations using multiple sophisticated measures and presenting coherent findings. Equally impressive is their collaborative skill development they work together seamlessly now, supporting each other's learning genuinely. The cooperative learning structures that felt awkward and effortful in Cycle I have become second nature."

Final reflection on the complete three-cycle intervention identified key factors contributing to success: the cyclical action research structure that allowed continuous refinement based on evidence, the systematic progression from simple to complex concepts with adequate time for consolidation, the use of varied cooperative structures matched to specific learning goals, the emphasis on conceptual understanding and interpretation rather than just procedural skill, the integration of visual representations with numerical measures, the use of authentic data and meaningful contexts, and the explicit attention to both mathematical content and collaborative skill development.

e. Cross-Cycle Analysis of Learning Gains

Examining trends across all three cycles provides insight into the trajectory of students' development and the cumulative impact of the cooperative learning intervention. Figure 1 (not shown) illustrates mean test scores across all assessment points: pre-test (M = 58.3), Cycle I post-test (M = 68.9), Cycle II post-test (M = 76.4), and Cycle III post-test (M = 82.7). The consistent upward progression with gains at each transition point indicates sustained improvement rather than a one-time boost followed by plateau.

The rate of improvement varied across cycles, with the largest absolute gain occurring between pre-test and Cycle I (10.6 points), followed by Cycle II to Cycle III (6.3 points), with the smallest gain between Cycle I and Cycle II (7.5 points). However, effect sizes remained large throughout: pre-test to Cycle I (d = 0.85), Cycle I to Cycle II (d = 0.72), Cycle II to Cycle III (d = 0.68). These sustained large effects suggest that the intervention continued to produce meaningful learning even as absolute scores increased and further improvement became more challenging.

Classical mastery rates showed similar progressive improvement: 27.8% at pre-test, 61.1% after Cycle I (gain of 33.3 percentage points), 72.2% after Cycle II (gain of 11.1 percentage points), and 83.3% after Cycle III (gain of 11.1 percentage points). The pattern of larger initial gains followed by continued but more modest gains is typical of learning trajectories and reflects the reality that foundational improvements are often easier to achieve than refined mastery of complex skills.

Individual student trajectories reveal differential patterns of growth that provide additional insight. Of the 18 students, 14 (77.8%) showed consistent improvement across all three transition points (pre-test to Cycle I, Cycle I to Cycle II, Cycle II to Cycle III). Three students (16.7%) showed strong improvement in two transitions but slight decline or plateau in one transition, suggesting that their learning was generally positive but not entirely linear. Only one student (5.6%) showed a more erratic pattern with substantial improvement in Cycles I and III but slight decline in Cycle II, possibly due to factors outside the intervention such as absence during key lessons or personal circumstances affecting engagement.

Examination of which students made the greatest gains reveals interesting patterns. The five students with lowest pre-test scores (range: 40-50) made an average gain of 28.2 points by the final post-test (average final score: 75.2), just reaching mastery threshold. The middle eight students (pre-test range: 55-65) made an average gain of 23.6 points (average final score: 82.8), exceeding mastery threshold comfortably. The five highest-performing students on pre-test (range: 70-80) made an average gain of 15.8 points (average final score: 88.4), achieving excellence level. This pattern suggests that the intervention benefited students across the achievement spectrum, though highest initial achievers had less room for improvement given ceiling effects.

The narrowing of score dispersion across cycles (SD decreased from 12.7 at pre-test to 9.1 at Cycle III post-test) indicates that lower-achieving students made proportionally greater gains, reducing achievement gaps. This finding is consistent with cooperative learning research showing that heterogeneous groups benefit lower achievers through peer scaffolding while still supporting continued growth for higher achievers through the cognitive benefits of explaining to others

f. Development of Cooperative Learning Behaviors

Beyond academic achievement, an important outcome of this intervention was students' development of collaborative skills essential for productive group work. Systematic observation across cycles documented progressive improvement in multiple dimensions of cooperative behavior.

Equitable participation defined as all group members contributing substantively to discussions and tasks showed marked improvement. In early Cycle I lessons, only 33% of groups demonstrated equitable participation patterns, with most groups showing dominance by one or two members. By late Cycle I, this increased to 67%. In Cycle II, equitable participation was observed in 83% of groups consistently. By Cycle III, all six groups (100%) regularly demonstrated balanced participation, with no groups showing persistent patterns of marginalization or dominance.

Active listening behaviors including eye contact with speakers, nodding, asking clarifying questions, and avoiding interruptions increased from 56% of students demonstrating these behaviors in Cycle I to 78% in Cycle II and 94% in Cycle III. By the end of the intervention, active listening had become a normative behavior in the classroom rather than an exception.

Quality of explanations provided during peer tutoring moments evolved substantially. In Cycle I, most explanations consisted of simply showing peers what to do ("First you find the biggest number, then the smallest number, then subtract"). By Cycle II, explanations increasingly included why elements ("You subtract them because range measures how far apart the most extreme values are"). By Cycle III, explanations often included connections to concepts ("Range tells you the total spread, but IQR is better when you have outliers because it focuses on where most of the data are").

Help-seeking behaviors also shifted in productive ways. Initially, when groups encountered difficulties, they typically called the teacher immediately. As cooperative skills developed, groups increasingly attempted to resolve difficulties themselves first, with one member explaining to others or the group collectively reasoning through confusion. Teacher intervention requests shifted from frequent calls for help with straightforward procedural questions to less frequent requests for clarification of genuinely complex conceptual issues.

Constructive conflict resolution emerged as a particularly important cooperative skill. In early cycles, disagreements sometimes led to unproductive arguments or one member simply insisting they were correct without justification. By Cycle III, disagreements typically prompted groups to return to evidence (data, definitions, worked examples) to adjudicate between competing perspectives. This evidence-based approach to resolving intellectual conflicts represents a sophisticated metacognitive strategy valuable beyond mathematics learning.

Teacher observation notes documented changing classroom culture across the intervention. Initial entries described the classroom as "teacher-dependent" with students "waiting to be told what to do." Middle entries noted "growing independence" and "students helping each other without prompting." Final entries characterized the classroom as "a genuine learning community where students take responsibility for their own and each other's learning." This cultural shift from individual, teacher-centered learning to collaborative, student-centered learning represents a foundational change in classroom dynamics.

2. Discussion

The findings from this three-cycle classroom action research study provide substantial evidence that cooperative mathematics learning can effectively enhance elementary students' statistical thinking skills related to data dispersion. This section discusses these findings in relation to existing literature, explores potential mechanisms explaining observed improvements, considers limitations and alternative explanations, and draws implications for educational practice and future research.

a. Effectiveness of Cooperative Learning for Statistical Thinking

The primary finding that cooperative learning produced large and sustained improvements in students' understanding of data dispersion aligns with and extends existing research on cooperative learning in mathematics education. The overall effect size from pre-test to final post-test ($d = 2.21$) substantially exceeds typical effect sizes reported in meta-analyses of cooperative learning interventions (Capar & Tarim, 2015, reported average $d = 0.59$; Johnson, Johnson, & Stanne, 2000, reported $d = 0.66$ for mathematics specifically). Several factors may explain the particularly strong effects observed in this study.

First, the extended duration and iterative refinement enabled by the action research design may have allowed the intervention to accumulate effects over time more successfully than shorter-term studies. Many experimental studies of cooperative learning span only a few weeks, potentially insufficient time for students to develop the collaborative skills and trust necessary for cooperative structures to function optimally. In contrast, this study's three-cycle structure spanning an entire semester gave students extended experience with cooperative learning, allowing both cooperative skills and mathematical understanding to deepen progressively.

Second, the specific content focus on statistical thinking about dispersion may have been particularly well-suited to cooperative learning approaches. As Garfield and Ben-Zvi (2008) argue, statistical reasoning inherently involves interpretation, argumentation, and consideration of multiple perspectives cognitive activities that are naturally supported by collaborative discussion. When students work alone on statistical tasks, they may calculate correctly without engaging deeply with what measures mean or why they matter. Cooperative structures that require articulation and justification of reasoning push students beyond procedural execution toward genuine conceptual engagement.

Third, the systematic attention to both mathematical content and collaborative process may have created synergies. Rather than treating cooperative learning merely as a grouping arrangement, the intervention explicitly taught collaborative skills, structured interactions to promote productive discourse, and monitored group processes to intervene when necessary. This dual focus ensured that group work actually produced the kinds of interactions elaborated explanations, cognitive conflict and resolution, perspective-takingthat learning theory suggests should promote conceptual development (Webb, 1991; Gillies, 2016).

The progressive improvement across cycles, rather than a single large initial gain followed by plateau, suggests that the learning supported by cooperative structures was cumulative and generative rather than simply a matter of increased time-on-task or novelty effects. Each cycle built on previous foundations, with students developing increasingly sophisticated understanding as they encountered more complex content and applied previously learned concepts in new contexts. This pattern supports sociocultural learning theories that emphasize learning as progressive participation in increasingly complex practices within communities of learners (Lave & Wenger, 1991).

b. Mechanisms of Learning in Cooperative Groups

While the quantitative data demonstrate that learning occurred, qualitative data from observations and recordings provide insight into how cooperative structures supported statistical thinking development. Several mechanisms appear particularly important.

Elaborated Explanation: Consistent with Webb's (1991) research, instances where students provided detailed explanations to peers were associated with learning for both explainers and recipients. The cognitive activity of organizing one's understanding sufficiently to explain it to another person requires deeper processing than simply solving problems individually. In this study, cooperative structures that created natural occasions for explanation (such as Jigsaw, where students taught teammates about their "expert" topics) appeared especially effective. Students who frequently explained concepts to peers tended to show larger learning gains, supporting the hypothesis that explaining is a powerful learning mechanism.

Cognitive Conflict and Resolution: Piagetian theory emphasizes the role of disequilibrium in prompting cognitive development (Piaget, 1985). Cooperative learning naturally generates cognitive conflict when group members hold different initial understandings or arrive at different solutions. Rather than being problematic, these moments of disagreement when well-managed create opportunities for deeper thinking as students must articulate their reasoning, consider alternatives, and reconcile differences. Observations documented numerous instances where disagreements prompted groups to return to definitions, examine worked examples, or carefully analyze data, ultimately leading to improved understanding for all members.

Multiple Representations and Perspectives: Statistical concepts like dispersion can be represented numerically (through calculated measures), visually (through graphs), and verbally (through descriptions and interpretations). Working in groups exposed students to multiple ways of thinking about the same concepts, as different group members spontaneously approached problems through different representational modes. This representational flexibility moving fluidly between numbers, graphs, and languageis characteristic of statistical expertise (Wild & Pfannkuch, 1999) and appeared to develop more readily in cooperative than in individual contexts.

Scaffolding within the Zone of Proximal Development: Vygotsky's (1978) concept of the ZPD suggests that learning is maximized when students work on tasks slightly beyond their independent capability with support from more knowledgeable others. Heterogeneous cooperative groups naturally create ZPD conditions, with higher-achieving students providing scaffolding for lower-achieving peers. Observations confirmed that peer scaffolding frequently occurred, with more capable students not simply giving answers but rather asking guiding questions, providing hints, and modeling problem-solving strategies that helped struggling students progress.

Development of Statistical Discourse and Identity: Over time, cooperative structures appeared to support students' developing identities as competent mathematical thinkers and their fluency with statistical language. In classrooms where the teacher is the primary or only source of mathematical validation, students may remain passive recipients of knowledge. In cooperative classrooms, students become authorities who explain, justify, and evaluate mathematical claims. This shift in epistemic role appeared to increase students' confidence and willingness to engage with challenging concepts. By Cycle III, students routinely used sophisticated statistical vocabulary in natural conversation, suggesting they had internalized both the language and concepts of statistical thinking.

c. Alignment with Statistical Education Research

The findings align well with recommendations from statistics education research regarding effective approaches for teaching variability and developing statistical thinking. Bakker and

Gravemeijer (2004) argue that understanding distribution and variability requires moving beyond isolated calculations to comparative reasoning where students analyze how different datasets or distributions differ and what those differences signify. The intervention's emphasis on comparison tasks comparing distributions with same ranges but different patterns, comparing datasets using multiple measures, comparing contexts where different measures are more informative directly reflects this recommendation and appears to have been effective.

Makar and Rubin's (2009) framework for informal inferential reasoning emphasizes that statistical thinking involves making inferences that go beyond the data at hand, articulating uncertainty, and grounding claims in evidence from data. Cooperative learning activities in this study regularly engaged students in these practices. When groups presented findings from their statistical investigations in Cycle III, they made claims about patterns, supported claims with specific measures and visual evidence, and acknowledged limitations all components of informal inferential reasoning. The collaborative context appeared to naturally prompt these practices because students had to convince their peers, not just satisfy themselves, that conclusions were justified.

Garfield and Ben-Zvi's (2005) framework for teaching and assessing reasoning about variability emphasizes the importance of (a) recognizing that variability is everywhere in data, (b) measuring variability appropriately, (c) representing variability visually, and (d) recognizing the impact of sampling on variability. This intervention addressed the first three components explicitly (the fourth, sampling variability, was beyond elementary curriculum scope). The progressive development observed—from initial vague awareness of variability to sophisticated use of multiple measures and representations—mirrors the developmental trajectory Garfield and Ben-Zvi describe, suggesting that the instructional sequence was developmentally appropriate and effective.

The persistent challenge students faced in articulating why certain measures are preferable in particular contexts (reflected in lower performance on these highest-level items even on the final post-test) aligns with research showing that metacognitive understanding knowing when and why to use particular tools develops more slowly than procedural skill (Garfield & Ben-Zvi, 2008). Even professional statisticians report that developing judgment about which methods suit which situations requires extensive experience. The fact that elementary students in this study showed emerging competence in this area, even if not mastery, represents significant progress.

d. Impact on Different Learner Groups

An important equity consideration is whether cooperative learning benefits all students or primarily advantages certain groups. Data from this study suggest that the intervention was broadly beneficial, though patterns of benefit differed across initial achievement levels.

Lower-achieving students showed the largest absolute gains (average improvement of 28.2 points), bringing their mean performance from failing levels to just reaching mastery threshold. This substantial progress suggests that cooperative learning successfully provided scaffolding and support that these students needed but were not receiving in traditional instruction. Peer explanations appeared particularly valuable for these students, providing multiple exposures to concepts through different explanatory approaches and in more accessible language than formal teacher instruction alone might offer.

However, it is also notable that lower-achieving students, despite large gains, remained the most likely not to achieve mastery (3 of 5 lowest pre-test performers did not reach 75% on final post-test, though they scored 70-74%). This suggests that while cooperative learning substantially helped these students, additional targeted support such as small-group intervention focused on foundational concepts, one-on-one tutoring, or extended time on prerequisite skills might be necessary for all students to reach mastery levels in complex domains like statistical thinking.

Middle-achieving students showed strong gains (average 23.6 points) and the highest rate of reaching mastery (100% achieved ≥75% on final post-test). These students appeared to benefit from both receiving explanations from higher-achieving peers and providing explanations to lower-achieving peers, consistent with research showing that teaching others is a powerful learning mechanism (Roscoe & Chi, 2007).

Higher-achieving students showed smaller absolute gains (average 15.8 points) due to ceiling effects—their pre-test scores were already relatively high. However, qualitative data suggest these students still benefited substantially from cooperative learning in ways not fully captured by test scores. Observations revealed that higher-achievers deepened their conceptual understanding through the challenge of explaining to others, developed more sophisticated mathematical communication skills, and demonstrated leadership development. Their final test scores (average 88.4) placed them at advanced levels, suggesting they developed expertise beyond basic mastery.

No clear gender differences in learning gains were apparent in this small sample (males gained an average of 23.8 points, females 24.2 points), though the study was not powered to detect gender effects. Observations suggested that mixed-gender groups functioned well, with no consistent patterns of male dominance or female marginalization, possibly because of explicit attention to equitable participation as a instructional goal.

C. Conclusion

This three-cycle classroom action research study demonstrates that cooperative mathematics learning significantly enhances elementary students' statistical thinking skills on data dispersion. The intervention successfully transformed students from having limited procedural knowledge to developing sophisticated statistical reasoning capabilities, enabling them to not only calculate measures of spread but also interpret what these measures reveal about data distributions and select appropriate analytical tools for different contexts.

The study reveals that cooperative learning structures support statistical thinking development through several key mechanisms. Elaborated explanation requires students to articulate their understanding clearly, benefiting both explainers and recipients. Cognitive conflict arising from peer disagreements prompts deeper examination of evidence and reasoning. Multiple representational approaches numerical, visual, and verbal emerge naturally in group discussions, developing the flexibility characteristic of statistical expertise. Peer scaffolding within heterogeneous groups supports struggling learners while challenging higher achievers through teaching responsibilities.

Beyond academic achievement, the intervention fostered significant development in collaborative skills. Students evolved from teacher-dependent learners to active members of a genuine learning community, demonstrating equitable participation, active listening, constructive disagreement, and shared responsibility for learning. This transformation in classroom culture represents a fundamental shift from individual, teacher-centered instruction to collaborative, student-centered learning.

The action research methodology proved crucial for success, allowing continuous refinement based on evidence of student learning and engagement. The cyclical process of planning, implementing, observing, and reflecting enabled the intervention to address challenges systematicallysuch as unequal participation and difficulty articulating conceptual understanding through targeted adjustments in subsequent cycles. This iterative approach optimized effectiveness in ways that fixed experimental designs cannot achieve.

Several factors contributed to the intervention's success. The systematic progression from simple concepts to complex ones allowed cumulative understanding. Real, personally meaningful data increased engagement and made abstract concepts concrete. Explicit instruction in cooperative skills alongside mathematical content ensured productive interactions. Integration of visual representations with numerical calculations supported multiple learning pathways. Consistent emphasis on interpretation rather than mere calculation fostered genuine statistical thinking.

The study also highlights important implementation considerations. Cooperative learning requires substantial teacher skill in task design, group formation, and productive intervention. The intensive nature of implementation demands careful planning, ongoing observation, and explicit attention to both content and collaborative processes. While highly effective overall, cooperative learning should be part of a comprehensive instructional approach that includes targeted support for students with the greatest needs.

The findings have significant implications for educational practice. Teachers should integrate well-designed cooperative learning structures into mathematics instruction, particularly for conceptually complex content. Curriculum developers need materials that genuinely require collaboration, with explicit guidance for facilitating group processes. Teacher educators must prepare teachers through both coursework and practice-based experiences in implementing cooperative learning effectively. The study also demonstrates the value of practitioner-led action research for developing pedagogical innovations in authentic classroom contexts.

E. References

Bakker, A., & Gravemeijer, K. P. (2004). Learning to reason about distribution. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht: Springer Netherlands.

Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. *The challenge of developing statistical literacy, reasoning and thinking*, *66*, 3-15.

Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed-ability approach. *British Educational Research Journal*, *34*(2), 167-194.

Capar, G., & Tarim, K. (2015). Efficacy of the cooperative learning method on mathematics achievement and attitude: A meta-analysis research. *Educational Sciences: Theory & Practice*, *15*(2), 553-559.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, *10*(3).

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning*, *1*(1), 5-43.

Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. *The challenge of developing statistical literacy, reasoning, and thinking*, 375-396.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, *31*(3-4), 175-190.

Cochran-Smith, M., & Lytle, S. L. (2009). *Inquiry as stance: Practitioner research for the next generation*. Teachers College Press.

Cooper, L. L., & Shore, F. S. (2010). The effects of data and graph type on concepts and visualizations of variability. *Journal of Statistics Education*, *18*(2), 1-16.

Davidson, N., & Major, C. H. (2014). Boundary crossings: Cooperative learning, collaborative learning, and problem-based learning. *Journal on Excellence in College Teaching*, *25*(3-4), 7-55.

delMas, R. C., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55-82.

Driver, R., & Bell, B. (1986). Students' thinking and the learning of science: A constructivist view. *School Science Review*, *67*(240), 443-456.

English, L. D. (2012). Data modelling with first-grade students. *Educational Studies in Mathematics*, *81*(1), 15-30.

Ferrance, E. (2000). *Action research*. Northeast and Islands Regional Educational Laboratory at Brown University.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*(2), 124-158.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-25.

Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, *4*(1), 92-99.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer.

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education*, *41*(3), 39-54.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371-404). Information Age Publishing.

Hossain, A., & Tarmizi, R. A. (2013). Effects of cooperative learning on students' achievement and attitudes in secondary mathematics. *Procedia-Social and Behavioral Sciences*, *93*, 473-477.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, *38*(5), 365-379.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. University of Minnesota.

Kemmis, S., & McTaggart, R. (2005). Participatory action research: Communicative action and the public sphere. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 559-603). SAGE Publications.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Makar, K., & Confrey, J. (2005). "Variation-talk": Articulating meaning in statistics. *Statistics Education Research Journal*, *4*(1), 27-54.

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

Mertler, C. A. (2017). *Action research: Improving schools and empowering educators* (5th ed.). SAGE Publications.

OECD. (2019). *PISA 2018 assessment and analytical framework*. OECD Publishing.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. University of Chicago Press.

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201-226). Kluwer Academic Publishers.

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, *77*(4), 534-574.

Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, *134*(2), 223-246.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, *10*(3).

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

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14.

Slavin, R. E. (2015). Cooperative learning in elementary schools. *Education 3-13*, *43*(1), 5-14.

Stigler, J. W., & Hiebert, J. (2009). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Free Press.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Falmer Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

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

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*(5), 366-389.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-248.

Zakaria, E., Chin, L. C., & Daud, M. Y. (2010). The effects of cooperative learning on students' mathematics achievement and attitude towards mathematics. *Journal of Social Sciences*, *6*(2), 272-275