The Role of Composition, De-composition, and Comparison Activities in 3–4-Year-Old Children's Abstraction Processes in Shape Recognition

Ayşe Pınar Şen, B.A., M.A.

Department of Educational Studies

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy in Educational Studies

Faculty of Education, Brock University
St. Catharines, Ontario

Abstract

This mixed-methods study investigated the impact of the Froebel Gifts Early Geometry Program (FG-EGP) on the development of abstraction processes in geometric shape recognition among 3- to 4-year-old children. Grounded in cognitive theories of abstraction, the study aimed to determine if structured geometry instruction using manipulatives, comparative activities, and explicit vocabulary could support a shift from visual to property-based reasoning in early shape classification. A total of 93 children and nine early childhood educators from four licensed daycare centres in the Niagara Region participated in the study. Classrooms were assigned to either experimental or control groups, with further division based on intervention duration (4-week or 8-week). Quantitative data were collected through pre- and post-tests using the Geometric Shape Recognition Test (GSRT), assessing children's ability to recognize typical and atypical examples of circles, squares, rectangles, and triangles and to articulate geometric properties. Qualitative data were gathered through structured classroom observations and semi-structured teacher interviews, focusing on instructional activities, vocabulary use, and manipulative mediation.

Quantitative analyses using repeated measures ANOVA, ANCOVA, and non-parametric tests revealed that children in the experimental groups, particularly those in the 8-week intervention, demonstrated significantly greater gains in both total recognition scores and property-based verbal responses compared to control groups. These findings indicate that the FG-EGP supported the development of abstraction by promoting conceptually rich interactions with shapes and enhancing geometric vocabulary. Qualitative findings further illustrated how structured activities, teacher scaffolding, and the use of Froebel Gifts facilitated children's engagement with geometric properties and supported deeper mathematical reasoning.

The integration of quantitative and qualitative results provides compelling evidence that early geometry instruction, when thoughtfully designed and developmentally appropriate, can foster abstraction in preschool-aged children. This study offers implications for curriculum design, teacher training, and early mathematics education policy aimed at strengthening foundational cognitive skills in young learners.

Acknowledgements

This dissertation marks the culmination of a long and transformative journey that would not have been possible without the presence, support, and guidance of many remarkable people.

Above all, I am deeply grateful to God. Through moments of uncertainty, fatigue, and doubt, it was faith that carried me forward and reminded me of the purpose behind this path. Without this divine strength, none of this would have been possible.

This Ph.D. process has been more than an academic pursuit; it has been a profound journey of self-discovery and growth. It not only helped me become a more rigorous researcher and a true seeker of knowledge, understanding what scientific information really means, but also taught me how to find truth. Along the way, I came to realize what I am capable of. This realization, born from perseverance and hard work, is one of the greatest feelings I have ever known, something that cannot be fully described, photographed, or titled. The confidence that emerged from this process has revealed a potential I always believed I had but had never fully seen until now. I will carry this feeling with me always. I now know, through experience, that I am capable of achieving anything I dream of. Nothing can hold me back as long as I keep faith in God and carry that divine essence within me.

I owe my deepest love and gratitude to my parents, İsmet and Filiz, and my brother Ozan. You have stood by me with quiet strength, endless encouragement, and sacrificial love. Your loving presence and steady support gave me comfort and strength at every step, making this journey feel lighter and more meaningful. This achievement belongs to you as much as it does to me.

I am especially grateful to my supervisor, Dr. Mary Louise Vanderlee, whose insightful suggestions and consistently well-timed, critical questions helped clarify and strengthen the direction of my research. Her mentorship has been exceptional, steady, and generous, at times subtle, at times direct, always thoughtful, and attuned to what I needed in each stage of this journey. Her trust in my vision, combined with her steady guidance and encouragement, has been instrumental in shaping both this dissertation and my growth as a scholar.

I also wish to thank my committee members, Dr. Joyce Mgombelo and Dr. Alex Lawson. Their critical insights, support, and valuable feedback challenged me to think more deeply and refine my work. I am thankful to Dr. Mgombelo for her thoughtful provocations, which invited me into a more reflective and expansive understanding of mathematics education. I am thankful to Dr. Lawson, who continually encouraged me to think about what my results and findings truly mean, not only statistically but also conceptually, and to consider their broader significance within early childhood mathematics education.

To the teachers, children, and administrators who welcomed me into their classrooms and lives, thank you for your openness and generosity. This research would not exist without your participation.

Finally, to the friends and colleagues who listened, encouraged, and reminded me of my strength during difficult times, thank you. Your words and presence mattered more than you know. This dissertation is the product of many hands, hearts, and moments. I carry the imprint of each of you with gratitude.

Through this process, I have come to affirm more deeply what I have always intuitively believed: the reality created by God is living, continuous, and unfolding, and it appears ever-changing to a mind that seeks to understand it. Any knowledge we claim at a given moment is, at best, a snapshot of a passing instant, valuable yet transient, and my understanding of reality remains partial and developing. This journey has taught me that I can consciously participate in and respond to the reality God has created, engaging with it in ways that are dynamic and evolving. I now perceive myself more intentionally as a joyful warrior, not because challenges disappear, but because I understand that challenges are neither good nor bad; they simply are. They are given to me precisely because of my capacity to meet them in a joyful manner, transform through them, and continue forward with faith, courage, and purpose.

This is not an ending, but a continuation of the journey with renewed faith, strength, and joy in the unfolding reality I am enacting.

Table of Contents

List of Tables

Table **Page**

List of Figures

CHAPTER ONE: INTRODUCTION

This chapter introduces the study's focus on young children's abstraction in geometric shape recognition and the pedagogical conditions that support it. It begins by situating myself within mathematics education and early childhood education, and geographically within Canada, specifically Ontario and the Niagara Region, to provide contextual grounding for the study. It proceeds with situating geometry within Euclid's deductive framework and clarifies abstraction as a cognitive process of attending to defining properties, rather than surface features, as the conceptual background of the study. Drawing on empirical evidence that children frequently misclassify atypical and non-examples and rely on visual prototypes, the chapter articulates the problem the study addresses and introduces the Froebel Approach, with Froebel Gifts and structured comparison, composition, and de-composition tasks, as a response to the stated problem. The chapter then states the purpose, research questions, and hypotheses guiding the investigation, outlines the theoretical framework, and explains the study's significance for early childhood mathematics instruction and teacher development. Finally, it previews limitations and the organization of subsequent chapters.

Researcher Positionality and Context

My professional and scholarly work is situated at the intersection of early childhood education and mathematics education, with a particular focus on early geometry learning and abstraction processes. I completed my undergraduate degree in Early Childhood Education at Çukurova University and worked as a kindergarten teacher for three years in southern Turkey. During this time, I regularly used concrete materials and multiple instructional strategies in mathematics and geometry activities, which led me to question how different pedagogical approaches and manipulative-based designs influence young children's mathematical thinking.

Differences I observed between private and public-school instructional approaches, particularly in the use of materials, inquiry, and experiential learning shaped my early awareness that how mathematics is taught strongly affects how it is learned.

My own educational experiences further reinforced this perspective. I experienced both creative, material-rich mathematics instruction and highly formula-driven, teacher-centered instruction. Regaining confidence in mathematics through analogy-based and concrete instructional methods influenced my later teaching practice with young children, where I intentionally used manipulatives, metaphors, and structural representations to support conceptual understanding. These experiences led to my master's research, where I conducted a quasi-experimental quantitative study examining the effects of Froebel Gifts on young children's geometry skills, finding statistically significant gains across multiple shape and spatial sub-skills.

This line of inquiry continued into my doctoral work in Canada, where I am situated within mathematics education and early childhood education research in Ontario, specifically in the Niagara Region. My work aligns with Ontario's educational priorities in early years mathematics, including the Ontario Early Math Strategy and curriculum emphasis on spatial reasoning, multiple representations, manipulative use, and concept-based instruction. Through my doctoral research and teaching work with pre-service teachers in Ontario, I examine how structured, manipulative-based geometric activity design supports abstraction and concept formation in young learners. My cross-context experience, spanning Turkish and Ontario educational systems provides an interpretive lens for understanding how curriculum policy, instructional design, and material use shape early mathematics learning environments.

Conceptual Background of Study

Geometry is a complex hierarchical system consisting of relationships between abstract spatial-numerical variables that may or may not correspond to real-world situations. Its complexity arises from the hierarchical and interconnected structure of spatial-numerical knowledge (Haves & Ansari, 2020). In this study, the domain of geometry is framed through Euclid's perspective, as interpreted in Fitzpatrick's (2008) English translation of *The Elements*, as it is widely accepted in the elementary geometry curriculum (Clements, 2003; Usiskin, 1972).

According to Fitzpatrick's (2008) edited version of *The Elements*, Euclid presents plane geometry as a deductive system founded on five postulates from which all other geometrical propositions can be logically derived:

1. A straight line may be drawn between any two points.
2. Any terminated straight line may be extended indefinitely.
3. A circle may be drawn with any given point as the centre and any given radius.
4. All right angles are equal.
5. For any given point not on a given line, there is exactly one line through the point that does not meet the given line. (Russell, 1897)

In this framework, Euclidean geometry is built upon fundamental geometric entities such as the point and the line. As interpreted by Fitzpatrick (2008), Euclid defines a point as "that which has no part," meaning it is an indivisible location without breadth, length, or width. His first postulate implies that any two points determine a straight line, forming the basis for connecting discrete locations in space. The second basic concept, the line, is defined as a "breadthless length," indicating that it has only one dimension—length. Euclid's second postulate asserts that a straight line can be extended indefinitely; Fitzpatrick notes that Euclid's second postulate states that a straight line can be extended, and this line may or may not be straight, implying that curves are a form of a line.

When two lines intersect, they form an angle in which the degree (x) varies depending on the positions of the lines. When two curved or multiple straight or curved lines intersect, they form a plane, in other words, a two-dimensional shape. Euclidean geometry considers the concept of the plane as an evenly lying surface which consists of lines. Since a line consists of points, a plane can also be considered as consisting of points. From Euclid's perspective, a geometric shape can be defined as a space that is bounded (e.g., an infinite space does not have boundaries, therefore cannot be considered as a geometric shape). The boundaries of a line are the beginning and ending points, the boundaries of a plane are the edges, and the boundaries of a solid are the planes (Fitzpatrick, 2008).

From this perspective, we can define the geometric shapes using only several scalar variables such as the number of lines, planes, edges, degree of angles, and length of sides. In other words, two-dimensional shapes can be considered as bounded figures with variations in the number of edges, number, length and direction of lines, and degree of angles, whereas in the case of solids the variations of number and shape of planes are also taken into consideration.

Abstraction is fundamentally important in understanding this process because the abstraction, by nature, represents a compressed version of a more complex situation (Millidge, 2021; Willingham, 2009). Abstraction is the process of generating abstract concepts that represent the commonalities or similarities between objects, concepts, operations, or relations (Cetin & Dubinsky, 2017; Mitchelmore & White, 2007; Nurhasanah et al., 2017). In other words, abstraction is summarizing a more complex knowledge while ignoring certain kinds of information that are irrelevant to the object of thought or the problem (Gray & Tall, 2007; Kamii et al., 2001). An abstraction can be explained as zooming out to operate on the big picture and keeping only the related details in consideration (Martinez & Huang, 2011; Willingham, 2009). It is a compression of bigger data; in other words, chunking the information in working memory (Millidge, 2021; Willingham, 2009). It requires the cognitive ability to focus on the deep structures rather than paying attention to the surface features of the relationship between the objects, concepts, or operations (Dubinsky, 2002; Kelley, 1984; Willingham, 2009).

Providing examples of abstraction from various domains is crucial in establishing a broader definition for abstraction before proceeding to explain what abstraction means specifically in learning geometry. For instance, in physics, the problem of a uniform mass rolling on a plane can be represented by using the variables of mass, friction, position, velocity, and acceleration. In a real-world physics problem, this mass problem can be represented by each atom's mass and position, and velocity on a plane that causes a certain amount of friction. However, if the problem is about a uniform mass rolling on a frictionless or a smooth plane, then the representation of the problem requires ignoring irrelevant information such as friction. An abstraction takes into consideration specifically and only certain properties of the problem that are relevant to the solution. In the case of the mass and plane problem in physics, abstraction is the process of focusing on the properties such as the speed of motion, acceleration, and vector direction. However, properties such as the velocity of an atom in the plane, the temperature of the atmosphere surrounding the mass, or the figure of the mass are irrelevant to the specific problem at hand (Millidge, 2021). The purpose of thought is fundamentally important in abstraction. If one is thinking of birds, a blue jay and a cardinal are relevant and related under the category of birds. However, if the object of thought is the colour red, then a cardinal and a strawberry are more relevant and related than a blue jay and a cardinal.

In the case of recognizing geometric shapes, the process of abstraction can be described as considering only the variables related to the recognition problem at hand, such as number, length and direction of lines, degree of angles, and number and shape of plane surfaces instead of adding into consideration the practically infinite amount of variables such as the colour or width of lines or texture of plane surfaces, which are irrelevant to the recognition of the shape. According to the literature (Kelley, 1984; Korkmaz, 2017; Mitchelmore & White, 2007; Nurhasanah et al., 2013; Scheiner & Pinto, 2014), shape differentiation and generating a dynamic prototype representing a group of shapes accurately and comprehensively are two very important indicators of young children's abstraction processes in learning geometry. However, generally, the problems in recognizing geometric shapes in the early childhood period are related to children's tendency to consider the figure's overall appearance and match it with a limited and/or inaccurate visual prototype (Arnas & Aslan, 2010; Clements & Sarama, 2000; Gecu-Parmaksiz & Delialioglu, 2019; Hallowell et al., 2015). The most typical indication of a lack of geometric knowledge is that children have limited imagery of different geometric shapes and their variances such as rhombus, parallelogram, isosceles triangles, right-angled triangles, obtuse triangles, and right-angled isosceles triangles (Charlesworth, 2005; Gecu-Parmaksiz & Delialioglu, 2019). Having established the theoretical foundations of geometry and the role of abstraction as a cognitive process, the following section turns to the empirical context in which these concepts manifest, focusing on the challenges children face in recognizing and classifying geometric shapes.

Background of the Problem

Early geometry learning often reveals a persistent gap between what children could notice, that is defining properties such as edges, angles, vertices, and faces, and what they actually attend to—that is, overall visual resemblance and familiar prototypes. Across studies, young learners reliably identify prototypical shapes (e.g., circles, squares, equilateral triangles) yet struggle when shapes vary in orientation, ratio, skewness, or dimensional context (Aslan & Arnas, 2007b; Clements et al., 1999; Halat & Yeşil Dağlı, 2016; Verdine et al., 2019; Yeşil Dağlı & Halat, 2016). These patterns signal not merely performance errors but also a deeper difficulty with abstraction where children frequently privilege surface features over property-based definitions, leading to under-inclusive categorization and misclassification of atypical and non-examples. This empirical landscape establishes the core problem addressed in the present study, which is that the opportunities to form robust, property-centred concepts of shape are limited in typical early childhood settings.

Converging evidence suggests that these learner-level patterns are usually shaped by instructional conditions (Arnas & Aslan, 2010; Clements & Sarama, 2000). In many classrooms, geometry receives comparatively less time and is presented with a narrow range of typical exemplars, few deliberate comparisons, and limited discourse that foregrounds defining attributes (Oughton, 2023; Ponte et al., 2023). Teacher preparation in geometry, particularly confidence with property-based definitions and class–subclass relations, affects the selection of tasks, the use of manipulatives, and the kinds of explanations that children hear. Consequently, children's opportunities to encounter, compare, and reason about atypical and non-examples are constrained. The subsections that follow therefore synthesize the empirical patterns in children's shape recognition and examine instructional drivers (teacher knowledge and opportunity-to-learn factors) that help explain why prototype-driven reasoning persists and where targeted intervention can be most effective. The next section presents empirical evidence illustrating how these difficulties appear in young children's actual shape recognition performance.

Empirical Patterns in Children's Shape Recognition

Research shows that 92% of 5-year-old and 95% of 6-year-old children recognize and correctly classify circles (Arnas & Aslan, 2010). However, the percentage reduces in 3- and 4-year-old children, as only the 77% of 3 years old and 85% of 4 years old children recognized and categorized circles correctly. Meanwhile, 78% of 5-year-old and 84% of 6-year-old children recognize and classify squares correctly, but a small proportion of them put non-square rhombuses into the square category (Arnas & Aslan, 2010). Only 49% of the 3-year-old children recognized squares correctly, however 4-year-old children scored better than the 5-year-old children in recognizing squares as 81% of them categorized them correctly. Although 81% of 5-year-old and 89% of 6-year-old children recognized the rectangles and categorized them correctly, children have difficulty in recognizing rectangles occasionally since they consider all parallelogram shapes with two parallel edges and square-like shapes as rectangles (Arnas & Aslan, 2010). The percentage reduced in younger children, as 53% of 3-year-old and 58% of 4-year-old children categorized rectangles correctly. Children were even less accurate in recognizing triangles as only 67% of 5-year-old and 72% of 6-year-old children were successful in classifying triangles. However, 4-year-old children were more successful than the 5-year-old children, as 70% of 4-year-old children categorized triangles correctly, while 62% of 5-year-old children categorized the triangles correctly. Children rejected triangles claiming that it is "too pointy," "too flat," or "the top is not in the middle" (Arnas & Aslan, 2010). In another study, the most common triangle that they recognized was the equilateral triangle (Charlesworth, 2005).

Other studies investigating the shape categorization process of 3- to 6-year-old children (Arnas & Aslan, 2010; Clements & Sarama, 2011; Kesicioğlu et al., 2011; Posnansky & Neumann, 1976; Sinclair & Moss; 2012) show that young children are more familiar with typical geometric shapes, but they are less successful in recognizing atypical geometric shapes with different sizes, orientation or kurtosis, and often categorize by considering the visual characteristics of the shapes instead of property attributes. Kurtosis is a measure of the geometric shape of a curve, specifically how tall and narrow the peak is and how heavy or thin the tails are. A high-kurtosis curve looks more sharply peaked with thicker tails, while a low-kurtosis curve looks flatter and more spread out.

Gecu-Parmaksiz and Delialioglu (2019) compared the effects of virtual and physical manipulatives on 5- to 6-year-old preschool children's understanding of geometric shapes through a quasi-experimental research design. Their study collected data from children's pre and post-test using the Geometric Shape Recognition Task instrument. The results of their study showed that virtual or physical manipulatives were not significantly different in facilitating the recognition of the circle. Virtual manipulatives were found to be more effective in facilitating the recognition of triangles, squares, and rectangles. However, the most important indication of this study related to my research is that the children who used both virtual and physical manipulatives had difficulty in categorizing geometric shapes when their attributes such as size, position, direction, orientation, ratio, or skewness were changed. This means that the manipulatives either virtual or physical are not sufficient by themselves to facilitate children's shape recognition skills.

Another study that supports these findings that the children have difficulty in categorizing geometric shapes when their attributes such as size, position, direction, orientation, ratio, or skewness were changed is by Hallowell and colleagues (2015). This study indicates that first-grade children overestimate the significance of the triangular vertices (pointiness) and scaling demands caused errors in shape classification, and when a projected curvature was involved, children also had difficulty in translating lines in two-dimensional diagrams into three-dimensional diagrams. For example, when children had the task to match a 2D flat paper triangle to a flat triangle shape printed on a paper, they were successful. However, they had difficulty to match the 2D flat paper triangle to a pyramid by ignoring the square base, or to a cone by ignoring the circular base (Hallowell et al., 2015). These results showed that children's mental imagery and experiences consist of typical examples and non-challenging experiences where non-examples of shapes are not included in the learning process. Another study that supports these suggestions was conducted by Arnas and Aslan in 2010, which indicated that kindergarten children were successful in recognizing typical examples of shapes, but when there were atypical features such as a different size, an uncommon orientation, or skewness in terms of ratio, children were not successful in recognizing shapes with accurate classification. This might be due to the lack of property considerations about shapes since children initially focus on overall visual aspects of shapes which makes it harder to recognize the exact shape of a figure. This problem may also be related to teachers' limited knowledge of geometry and their self-efficacy in teaching geometry. In the following section, the problems in the professional development of in-service teachers are elaborated in detail associated with the problems in children's shape recognition. These consistent patterns in children's geometric reasoning point toward a deeper instructional issue, how teacher knowledge, preparation, and opportunities to teach geometry shape children's experiences and conceptual growth.

Instructional Drivers: Teacher Professional Development and Opportunity to Learn

One of the topics that the prospective teachers claimed to have been prepared least to teach is geometry (Jones et al., 2002). Research indicates that most pre-service teachers recognized and categorized geometric shapes based on their overall physical resemblance to visual prototypes and do not consider the definitive properties of shapes (Clements, 2003; Clements & Sarama, 2011; Fujita & Jones, 2006). Although early childhood educators do not necessarily teach the relations between classes of shapes, they need to know considerably more so they may adequately assess and teach children effectively. For instance, if a teacher asks the children to find the rectangles around the class and rejects a child's choice of square, the teacher would have to unteach the inaccurate rationale of categorization. This is the first reason that the professional development of in-service teachers in the domain of geometry is crucially important.

A study that surveyed 400 teachers regarding the math activities that they spend the most time on revealed that 67% of early childhood teachers focus on counting; 60% focus on sorting; 51% of them focus on numeral recognition; 46% on patterning; 34% on number concepts; 32% on spatial relations; 16% on making shapes; and 14% on measuring (Sarama & DiBiase, 2004). This study manifests another reason why the professional development of in-service teachers in the domain of geometry is crucially important as the teachers spent the least amount of time on geometry activities among all mathematics teachings. This can create a learning environment that limits children's exposure to atypical shapes, and limits their opportunities for discussion and experience attending to identifying typical, atypical, and non-examples of shapes. Taken together, these studies highlight that children's struggles with abstraction may be rooted in their abilities but rather in the limitations of instructional design, content exposure, and teacher preparation, leading directly to the problem this study seeks to address.

Statement of the Problem

Previous research demonstrates that the barriers to young children's development of abstract thinking supporting geometric shape recognition are:

1. Lack of exposure to atypical shapes (Arnas & Aslan, 2007a; 2007b; Aslan & Arnas, 2010; Clements et al., 1999; Hannibal & Clements, 2010; Satlow & Newcombe, 1998).
2. Lack of discussion of typical, atypical, and non-examples of shapes (Christie & Gentner, 2010; Gecu-Parmaksiz & Delialioglu, 2019; Loewenstein & Gentner, 2001; Kotovsky & Gentner, 1996; Yuan et al., 2017).

3. Lack of experience attending to the identifying features of typical, atypical, and non-examples of shapes (Aslan & Arnas, 2010; Clements et al., 2018; Hallowell et al., 2015; Kamii et al., 2001; Sarama & Clements, 2004; Verdine et al., 2017; Verdine et al., 2016; 2014).

These barriers cause certain difficulties in children's geometry learning such as (a) limited recognition and under-inclusive categorization of shapes (Arnas & Aslan, 2007b, 2010) and (b) visual apprehension of shapes without paying attention to definitive features (Clements & Sarama, 2000). Kindergarten children recognize basic typical shapes such as squares, equilateral triangles, rectangles, or circles. However, they have difficulties recognizing atypical shapes such as right-angled isosceles triangle, trapezoid, rhombus, hexagon, or pentagon (Arnas & Aslan, 2007a; 2007b; 2010; Gecu-Parmaksiz & Delialioglu, 2019). When children categorize shapes, they tend to exclude the shapes with skewness or kurtosis, and/or situated in an atypical orientation (Arnas & Aslan, 2010; Clements et al., 2018; Gecu-Parmaksiz & Delialioglu, 2019; Hallowell et al., 2015). Moreover, children recognize and define the shapes visually without paying attention to the definitive properties of shapes (Arnas & Aslan, 2010; Clements & Sarama, 2000; Kalénine et al., 2011). The problem with the visual apprehension of the shape is that it limits the understanding of the shape which also limits the children's definition of the shapes since they do not consider the components.

The barriers to children's acquisition of abstract thinking in geometry stem from the difficulties associated with decontextualized problems (Walker et al., 2018). Studies (Arnas & Aslan, 2007a; Clements et al., 2018) indicate that this problem is related to the content and quality of the educational materials and instruction provided to children in preschool mathematics activities, rather than children's lack of ability or cognitive maturity. Arnas and

Aslan (2007) revealed that a majority of geometry activity books and videos published for kindergarten children fail to introduce the geometry content adequately since they involve only typical geometric shapes. Furthermore, their study identified that the geometry content is usually presented in isolation without establishing a contextualized understanding. These studies point out that abstraction may be difficult for young children unless effective instruction that engages children in comparison, composition, and de-composition activities, the opportunity for reflection and discussion, and variety-rich geometry content and materials are provided (Cetin & Dubinsky, 2017; Dubinsky, 2002; Mitchelmore & White, 2007). Understanding these barriers clarifies the need for targeted instructional approaches that can bridge the gap between children's visual and abstract understanding of shapes, providing the rationale for this study's intervention design.

Rationale

Studies (Clements et al., 2018; Gecu-Parmaksiz & Delialioglu, 2019) indicate that children can learn richer concepts about shape if the education includes:

1. a variety rich geometry content (Gecu-Parmaksiz & Delialioglu, 2019);
2. a broad array of geometric tasks (Clements et al., 2018; Gecu-Parmaksiz & Delialioglu, 2019);
3. conversations with the learners about the features of shapes in detail (Clements et al., 2018); and
4. structural alignment among the concepts and operations (Christie & Gentner, 2010).

These studies suggest introducing children to a wider variety of shape classes including atypical examples and non-examples can refine the concept formation for an accurate abstraction. Moreover, engaging children in a broad array of geometric tasks such as comparison, composition and de-composition, or area and/or angle measurement activities that gradually transition from sensory concrete to integrated concrete thinking can be effective to facilitate abstract thinking in geometry. In sensory concrete mode of knowing, the child makes sense about a concept or procedure through sensory material. Integrated-Concrete thinking mode is stronger in terms of abstraction, because it refers to the combination of separate sensory concrete ideas in an interconnected knowledge network. Children with *sensory-concrete* knowledge require sensory material to comprehend a concept or procedure (Clements, 2000). For instance, the majority of children cannot solve problems involving larger numbers without the aid of concrete objects until age 5.5 (Sarama & Clements, 2016). *Integrated-concrete* knowledge is knowledge that is interconnected in specified ways. The strength of integrated-concrete thinking derives from the combination of distinct ideas into a network of interconnected knowledge. For children with this type of interconnected knowledge, knowledge of physical objects, actions performed on them, and symbolic representations are all interconnected within a solid mental structure. Children begin to learn through sensory concrete activities, and gradually they connect the concepts they learn to other ideas and situations and make meaningful inferences (Sarama & Clements, 2016). The conversations between the learners and the teacher about the attributes and definitive properties of shapes in detail are fundamentally important for their abstract thinking. Christie and Gentner (2010) suggest that structural alignment is crucial for relational abstraction, emphasizing the significance of alignment among the concepts and operations in developing abstraction skills. They mention that a sufficient set of sequential examples enable young children to learn new relational-spatial patterns and such learning is a result of the learner’s opportunity to align the sequential representations. As a summary, these studies suggest that an effective geometry instruction in early childhood education must include a variety of shape classes and tasks that are introduced from simple to complex, from sensory concrete to integrated concrete, allowing children to expand their visual mental imagery, vocabulary and experience about geometric shapes and promoting conversations between the children about their geometry knowledge. In light of this evidence, the present study introduces the Froebel Approach and Froebel Gifts as a theoretically grounded and empirically testable framework to enhance young children's abstraction processes in early geometry learning.

Purpose of the Study

The purpose of this research is to test the efficacy of the Froebel Approach (instructional methods) and Froebel Gifts (materials) through an intervention designed to support 3- to 4-year-old children's abstract thinking in geometry. Froebel Gifts (FG) introduce a rich variety of typical and atypical geometric shapes in a structural alignment of concepts from three to two-dimensional, from surface to line, from line to point which can facilitate children's relational abstraction (Christie & Gentner, 2010; Manning, 2005; Reinhold et al., 2017; Tovey, 2016; von Marenholtz-Bülow, 2007). The numerous near-comparison, far-comparison opportunities between the typical and atypical examples of geometric shapes that Froebel Gifts Early Geometry Program (FG-EGP) provides are expected to direct children's focus on the similarities and differences between the defining properties of shapes by attending to their edges, vertices, and surfaces.

The following points sum up the potential for overcoming the aforementioned barriers to children's abstraction processes in learning geometry that the Froebel Approach and Gifts may offer:

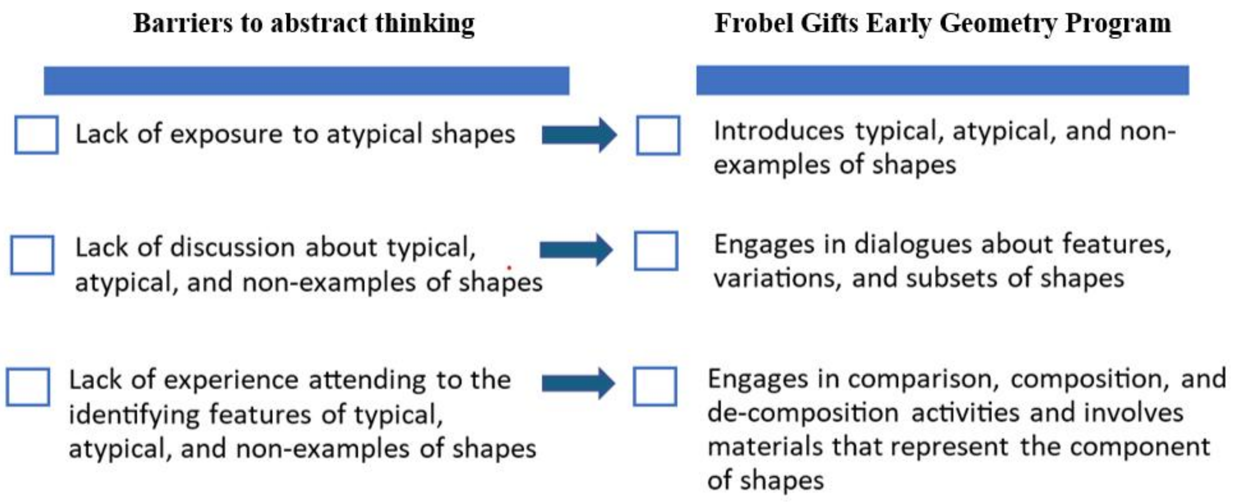
1. Lack of exposure to atypical shapes: Froebel Gifts include typical and atypical shapes and materials such as rods, beads, and tablets to form non-examples of shapes.

2. Lack of discussion about typical, atypical, and non-examples of shapes: Froebel's approach guides the teaching and learning process through conversation between the learners and teachers about the comparison of properties and characteristics of shapes.
3. Lack of experience attending to the identifying features of typical, atypical, and non-examples of shapes: Froebel Gifts include materials that represent the components of shapes such as sides, vertices, and plane surfaces and offer opportunities for composition and de-composition of three- and two-dimensional shapes.

Figure 1 illustrates how the Froebel Gifts Ealy Geometry Program addresses key barriers to children's abstract thinking in geometry by transforming each challenge into targeted instructional strategies that promote conceptual understanding through structured, hands-on activities.

Figure 1

Froebel Gifts as a Mediator to Overcome Barriers in Abstract Thinking in Geometry



Barriers to abstract thinking		**Frobel Gifts Early Geometry Program**
☐ Lack of exposure to atypical shapes	→	☐ Introduces typical, atypical, and non-examples of shapes
☐ Lack of discussion about typical, atypical, and non-examples of shapes	→	☐ Engages in dialogues about features, variations, and subsets of shapes
☐ Lack of experience attending to the identifying features of typical, atypical, and non-examples of shapes	→	☐ Engages in comparison, composition, and de-composition activities and involves materials that represent the component of shapes

Froebel Gifts have the potential of overcoming the noted barriers to abstract thinking in geometry as a mediator in geometry education. As children learn to compare, compose, and decompose geometric shapes through the instructions of the teacher within the scope of FG-EGP they usually use definitive properties (e.g., four sides, three corners, narrow, wide, or right-angled, parallel, etc.) when describing a shape. If the children use more visual rationale (e.g., it is rectangle because it looks like a door, or it must be a square, because it looks like a box, etc.), it implies that they are considering the shape's overall visual figure and matching it with a prototype (Halat &Yeşil Dağli, 2016; Jones, 1998; Ness & Farenga, 2016). If children use more definitive properties (e.g., it is a triangle because it has three straight sides and three corners, or it must be a rectangle because it has two parallel sides longer than the other two parallel sides, etc.), it means that the children can consider the relevant variables that define the shape at hand and ignore the irrelevant variables which is a sign of abstraction.

The instruction I developed emphasizes the similarities between geometric shapes' components such as sides, planes, and vertices and highlights the property similarity rather than an overall visual similarity. In this way, children can first disassemble the perceived three-dimensional objects into their components such as surfaces, edges, and vertices physically with the gifts and create mental concepts that represent these objects and operations. The gradual alignment of concepts in Froebel Gifts from three-dimensional to two-dimensional, from whole to part, from solid to plane, from plane to line, and from line to point may provide an effective strategy for concept formation. The comparison activities in FG-EGP are expected to help children distinguish the similarities and differences of geometric properties.

This study, in general, aims to investigate the effects of a geometry education program based on the Froebel Approach using Froebel Gifts on 3- to 4-year-old children's abstraction processes in learning geometry and on daycare teachers' experiences in teaching geometry. In specific, the purpose of this study is to explore the impact of a geometry education program that introduces children to typical, atypical, and non-examples of shapes and engages them in comparison, composition, and de-composition activities on teachers' teaching of abstract geometry concepts and facilitating children's abstract concept formation in geometry, specifically in shape recognition. Building upon the aims and theoretical grounding of the study, the following section outlines the specific research questions that guide the investigation, translating the study's broader purpose into measurable and observable lines of inquiry.

Research Questions

The first research question aims to determine whether children recognize shapes visually or recognize them through their definitive properties such as the number of edges, corners, and faces. An increase in children's property responses would indicate the development of abstraction skills in geometric shape recognition.

1. How do the differences in the geometry activities, manipulatives, and the vocabulary promoted by the teachers impact 3- to 4-year-old children's abstract thinking skills in shape recognition?

    1a) Does FG-EGP impact 3- to 4-year-old children's abstraction processes in geometric shape recognition?

        1a-1) If FG-EGP has an impact on children's abstraction processes in geometric shape recognition, how does it affect them?

    1-b) Does FG-EGP have an impact on children's property responses in shape recognition?

        1-b-1) If FG-EGP has an impact on children's property responses in shape recognition, does it affect them positively or negatively?

1-c) Does FG-EGP have an impact on children's usage of attributes and properties in defining shapes?

1-c-1) If FG-EGP has an impact on children's usage of attributes and properties in defining shapes, does it affect them positively or negatively?

The second research question aims to explore the daycare teachers' experiences with the manipulatives and instructions in teaching abstract concepts in geometry and the vocabulary they use. This question aims to collect data from the teachers to explain the possible changes in children's numerical scores that were obtained from the pre-test and post-test applications of Geometric Shape Recognition Test (Aslan & Arnas, 2007b).

2. What kind of activities, vocabulary, and manipulatives do the teachers use for children's engagement with the geometry activities before, during, and after the intervention?

   2a) What kind of activities do the teachers use for children's engagement with geometry concepts before, during, and after the intervention?

   2b) What kind of vocabulary do the teachers use for children's engagement with geometry concepts before, during, and after the intervention?

   2c) What kind of manipulatives do the teachers use for children's engagement with geometry concepts before, during, and after the intervention?

In alignment with these research questions, the subsequent section articulates the hypotheses that operationalize the study's expectations regarding how the Froebel Gifts Early Geometry Program (FG-EGP) influences children's abstraction and teachers' instructional practices.

Hypotheses

I hypothesized that the comparison, composition, and de-composition activities that FG-EGP provides can possibly influence children to focus on the property attributes of shapes such as side, vertices, angles, and edges of shapes and helps them to consider the shapes beyond their overall appearance which is an important aspect of the abstraction process. When children begin to think abstractly, they focus on the properties of shapes and start to consider a shape as a combination of its components such as angles, sides, vertices, and planes (Gray & Tall, 2007; Kato et al., 2002; Yee, 2019). Therefore, the increase in children's property responses to shape recognition questions were interpreted as a predictor of abstraction. The following hypotheses were tested:

- H1 – At post-test, children in the experimental group will have higher recognition accuracy than children in the control group, controlling for pre-test performance.
- H2 – At post-test, children in the experimental group will earn higher explanation scores (0 = "don't know", 1 = visual/prototypical, 2 = definitive/property-based) than children in the control group, controlling for pre-test explanation scores.
- H3 – Across observations and interviews, experimental classrooms will demonstrate and report higher use of manipulatives, comparison/composition/de-composition (CCD) activities, and property-focused vocabulary than controls; and, controlling for pre-test, greater observed/reported use will be associated with higher post-test recognition accuracy and a higher proportion of definitive (property-based) responses.

The hypotheses not only reflect the theoretical assumptions underpinning the study but also establish the foundation for its broader contribution to early geometry education. The next section therefore discusses the significance of this research within both scholarly and practical contexts, emphasizing its implications for pedagogy, curriculum design, and policy.

Significance of the Study

Much of the literature on the Froebel Approach and Froebel Gifts is historical or theoretical (Reinhold et al., 2017; von Marenholtz-Bülow, 2007; Zuckerman, 2006). Given their broad influence on the design of early mathematics manipulatives, there is a clear need for intervention-based evidence showing how these materials and accompanying instruction affect children’s geometric thinking (Zuckerman, 2006). Sen and Kotil (2022) reported that Froebel-mediated instruction improved recognition of 2D/3D shapes and their components (sides, vertices, plane surfaces) in 3–4-year-olds but did not test whether such gains reflect growth in abstraction. This study directly addresses that gap by examining whether comparison, composition, and de-composition activities with Froebel Gifts help children attend to definitive properties (e.g., sides, vertices, angles, faces) when identifying, recognizing, and classifying shapes.

Research on persistent difficulties in geometry learning points to overreliance on overall visual appearance and limited property-based reasoning (Arnas & Aslan, 2010; Ozerem, 2012). Scholars recommend that teachers use alternative definitions, rich property vocabulary, and varied representations to strengthen geometric arguments. Accordingly, this study integrates those recommendations into its design: children encounter variations of forming shapes (e.g., building a square with beads, rods, or planes) and variations of defining shapes (e.g., a square as two right-angled triangles; as four equal-length sides and four vertices; or as four right angles) to scaffold abstraction in shape recognition.

The work is also policy relevant. Ontario’s math strategy (Government of Ontario, 2020) emphasizes secure understanding of foundational concepts in mathematics and geometry for both teachers and students, and the *Early Math Strategy* (Ontario Ministry of Education, 2003) underscores the role of concrete, tactile materials. By focusing on 3–4-year-old children and in-service daycare teachers, this study responds to those priorities, as it targets core geometric constructs, edge, vertex, plane, and examines how structured, manipulative-rich instruction develops conceptual language and reasoning aligned with provincial goals.

Finally, the findings informed classroom practice and manipulative design. Children’s explanations revealed the concepts and vocabulary they actually use, clarifying how they perceive shapes and how their reasoning connects to other ideas. Teachers’ accounts of which activities and materials they need, and why, highlighted features of effective geometry tasks. Taken together, children’s outcome data and teachers’ perspectives yielded actionable guidance for curriculum design, instructional routines, and the next generation of manipulatives in early childhood geometry.

CHAPTER TWO: REVIEW OF THE LITERATURE

The goal of facilitating children’s abstract thinking in shape recognition is a complex issue in which the development occurs within children’s interaction with each other and their teachers, the manipulatives used in the learning environment, the instructions given by their teachers, the vocabulary promoted in the classroom, and the activities that the teachers do in the learning environment. Therefore, the problem involves multiple dimensions, including developmental aspects such as the development of abstract thinking and geometric concept formation in young children, as well as analytical aspects related to difficulties in children’s abstraction and geometry learning. It also includes instructional and professional dimensions, such as strategies for supporting abstraction in geometry, the role of manipulatives and language in abstract thinking, and early childhood teachers’ knowledge of geometric abstract concepts.

First of all, this study defines the concept of abstraction and describes the process of abstraction according to the existing theories in the literature. To establish a general understanding of the concept of abstraction, the abstraction theories of Piaget, Dienes, and Mitchelmore and White are described comparatively in detail. The literature review follows with the analysis of what abstraction means and what role it plays in children’s understanding of geometric shapes. Accordingly, the existing literature about the abstraction processes in geometry learning and shape recognition is analyzed to highlight the characteristics that an effective geometry program should include to allow and promote abstract thinking in young children. After establishing a broad and deep understanding of the development of abstract thinking in young children, this literature review continues with describing the theoretical framework that is used as a lens in this study. Fischbein’s figural concepts and Vinner’s concept image theories form the theoretical framework of this study. The theoretical framework illuminates the important indicators of young children's abstraction processes in learning geometry which sets the knowledge base for the analysis of the current issues in young abstract thinking children's shape recognition.

The existing literature regarding young children's abstract thinking in geometry learning and shape recognition is grouped into:

- Problems associated with children's abstract thinking in shape recognition and categorization
- Problems associated with in-service and pre-service teacher's knowledge and experiences in geometry teaching.

The existing literature regarding the problems associated with children's abstract thinking in shape recognition and categorization is grouped into three major themes: lack of exposure to atypical shapes; lack of discussion about typical, atypical, and non-examples of shapes; and lack of experience attending to the identifying features of typical, atypical, and non-examples of shapes. Finally, this literature review introduces Froebel Gifts and Froebel Approach as a solution to the aforementioned problems regarding children's abstract thinking in geometry learning. To systematically explore these developmental, instructional, and theoretical dimensions of abstraction in geometry, a structured literature search was conducted across major databases and scholarly sources.

Theories of the Development of Abstract Thinking

The process of abstraction is defined by several theorists such as Piaget, Skemp, Dienes, and Mitchelmore and White. The common definition of the term among these theories can be deduced in two essential processes. The process of similarity recognition is the first step of abstraction. For example, shapes X and Y must be distinguished from shapes Z, proving that X and Y are more similar. The second mental step entails forming a concept that would classify the observed similarities as a group. For instance, the shapes X and Y must go under category A, however Z does not, because of its distinct decisive features, and falls under a different class (Martinez & Huang, 2011). Recognizing commonalities is thus the first and most basic step in the abstraction process.

According to Piaget, there are two basic types of abstraction: reflective and empirical (as cited in Dubinsky, 2002). Piaget claimed that by concentrating on some characteristics of the objects and disregarding others, meaning is constructed. Since these features may be perceived empirically via our senses, this process is known as empirical abstraction (Mitchelmore & White, 2007; Nurhasanah et al., 2017; Nurhasanah et al., 2013). In the operation of reflective abstraction, when people conduct mental actions on mental conceptions, the mental operations themselves eventually turn into new objects of thought (Cetin & Dubinsky, 2017; Gray & Tall, 2007; Kamii et al., 2001; Kato et al., 2002). Reflective abstraction abstracts from conceptual knowledge as opposed to actual physical objects. Reflective abstraction is described by Piaget as a cognitive process or operation, stressing that it is not a fixed arrangement of sensory components. A mental operation that is created at one level of operation is abstracted from that level and used at a higher state in this dynamic process (Kato et al., 2002). Although Piaget describes the abstraction process in detail, his theory does not explicitly inform and associate with geometry education. However, Dienes explained abstraction as the cornerstone of mathematical education and the core of mathematics learning.

Dienes (1971; as cited in Gningue, 2006) not only described the abstraction processes but also informed about the learning and teaching experiences that form abstract concepts (as cited in Borasi, 1984). He proposed six steps to promote abstract thinking in mathematics:

1. Free play in situations which the concept itself is concretized.

2. Structured play in the situations to recognize and use basic properties and rules.
3. Comparison activities among the situations to recognize the common elements and abstract the concept itself.
4. Generating representations of the new concept to better identify and fix the prototype.
5. Developing a symbolism to express the properties discovered.
6. Formalization of the results. (Benedek, 2018; Borasi, 1984)

Dienes argues that the learning process must start with unstructured interaction with a number of examples in which the topic itself is concretized, followed by guided activities in which the basic properties and rules are identified and introduced. When a teacher is certain that the students understand the fundamental ideas and principles, they must gradually move on to comparison tasks among the examples in order to help the students identify the features that are shared by all of them and thereby abstract the idea itself (Borasi, 1984). When a child exhibits the first signs of abstraction, the teacher helps them find examples of the new concept so they may better understand it. Following this procedure, the teacher guides the students in creating meaningful symbols to represent the concepts that have been abstracted as well as their relationships. The formalization of the conclusions achieved in a rigorous and possibly axiomatic method is the completion of Dienes's learning and instruction phases (Benedek, 2018; Gningue, 2006; Sriraman & English, 2005).

These educational phases are experienced throughout the course of three time frames that Dienes referred to as the "cycles." With special focus on the first three stages that include the abstraction process, teachers must assist the students in effectively navigating through each step. The initial three exercises—free play, guided play, and comparative tasks—are related to Cycle 1. Children engage in explorations with concrete objects, start picking up fundamental concepts and rules, and create mental images of these concepts during these tasks. Children must see a concept in several contexts and comprehend its continuous characteristics during these operations. Children can play games after they have a basic understanding of some constant characteristics and patterns. For the learners to understand the mathematical concepts, this cycle must lead to abstraction.

The second cycle is about the transition from activities with manipulative materials to more abstract representations, such as graphics and pictures. Cycle 2 begins with the representation stage, after the child has observed the common elements in each example of the concept. For example, if the children can talk about the concept of angle, after learning the concept of line and vertices, they can also represent the angle's properties after they are shown an example of representation of the concept that embodies all the common elements found in each example, by the teacher. This could be measuring the angle with a goniometer and introducing the angle symbols in writing (Gningue, 2006). The third cycle begins when children recognize the commonality of the concept in the different representations and present the representations of the concept using accurate verbal and mathematical system of symbols (Gningue, 2016). My study focused on the first cycle as the aim is to investigate children's abstraction processes in shape recognition.

Dienes's philosophy sought to employ specific manipulatives to improve students' understanding of mathematical concepts (Dienes, 1971; as cited in Gningue, 2006). He thought that the necessity for abstraction and generalization was what made it difficult to grasp mathematical concepts. According to Dienes, abstraction is the process of categorising concepts, situations, and objects in diverse contexts so that each object or situation belongs to this class due to a specific attribute it holds. Abstraction would render all further qualities for that particular classification worthless (Borasi, 1984, p. 14). He explained the development of abstraction using his theories of mathematical learning and instruction. Dienes highlights abstraction as the cornerstone of mathematical education and the core of mathematics. Even though his emphasis on teaching mathematics with the principle of abstraction was criticized in terms of not being the most appropriate strategy when learning specific topics in mathematics such as whole numbers (as it is understood better with the notion of infinity rather than abstraction) (Borasi, 1984), it is critical in shape recognition.

Accordingly, Dienes stated four principles of mathematical learning as following:

1. Dynamic
2. Constructivity
3. Perceptual Variability (or Multiple Embodiment)
4. Mathematical Variability (Benedek, 2018; Gningue, 2016; Sriraman & English, 2005).

According to the Dynamic principle, transformations in one model (such as changing an isosceles triangle into a right-angled triangle) must match those in an isomorphic model (transforming a wide angled triangle to an equilateral triangle). In this context, isomorphism refers to a mapping between two structures of the same kind, such as two types of triangles, while maintaining the structure between the components of the sets. It can be reversed via an inverse mapping (Sriraman & English, 2005). This idea outlines three fundamental phases for thought generation (Sriraman & English, 2005). Children must engage in unstructured play in the initial stage in order to acquire informal knowledge about the task at hand. Children must then be exposed to a variety of activities that are structurally related to the targeted concepts. Children start to build mathematical concepts in the third and final stage, and they engage in activities that enable them to use these concepts in a variety of circumstances.

According to the Constructivity principle, children's development of mathematical concepts and their relationships among them is facilitated when they reflect on their thoughts about physical and mental operations on concrete objects. Children's constructive thinking develops before their analytical thinking, hence activities and games must be made so that children can construct before they analyse (Dienes, 1971; as cited in Gningue, 2006).

The Multiple Embodiment principle, also known as the Perceptual Variability principle, contends that concepts should be introduced to children through a variety of actions that take place in varied settings. Children could then recognize the isomorphic structures shared by various situations due to the fact that the children are given chances to abstract the structural (conceptual) mathematical similarities. Children would have the chance to connect structurally related problems to one another when the perceptual details of a situation or problem differ but share certain common structural traits (Sriraman & English, 2005). Instead of using just one representation of an idea, teachers must use a variety of tangible objects to maximize the conceptual learning in their students (Dienes, 1971; as cited in Gningue, 2006). For instance, teachers might direct students to utilize various materials (such as point, line segments, planes, three-dimensional objects, etc.) to construct basic geometric shapes so that students can see how certain properties or structures are shared by all of them (Gningue, 2006). For instance, by using two long and two short rods as line segments and beads as the vertices, they can create a rectangle. Alternately, students can construct the same shape entirely out of beads, which can help children realize that every line segment in geometry is really just a line of points.

The final principle, the Mathematical Variability principle, suggests that some "irrelevant traits" should be introduced to children. They must make use of them in relation to creating concept structures because doing so would draw attention to the constant characteristics in varied manipulations and strengthen the overarching mathematical idea in each one of them. For example, to improve students' understanding of the triangle, the teacher should vary as many irrelevant characteristics as possible (such as the length of the sides, the width of the angles, and the triangle's location) while maintaining consistency with the pertinent characteristic (closed shape with three straight sides) (Gningue, 2006).

Dienes's theory was criticized by Borasi (1984) regarding his focus on learning by abstraction in every subject in mathematics, because abstraction is not the principal cognitive process for subjects such as elementary calculus or number theory. For example, in elementary calculus, most of the concepts are developed mostly through potential applications in the real world, instead of by noticing properties common to diverse situations. However, abstraction is a fundamental process in learning elementary transformation geometry (Borasi, 1984).

Mitchelmore and White (2007) developed a theory of "Teaching for Abstraction" in which students:

- familiarize themselves with the structure of a variety of relevant contexts;
- recognize the similarities between these different contexts;
- reify the similarities to form a general concept;
- apply the concept in new situations.

In terms of shape recognition, this theory suggests that children must first get acquainted with a variety of a specific shape class (e.g., different types of quadliterals, or different types of triangles) to distinguish the similarities between the shapes which would direct children's attention specifically to the components or definitive properties of the shapes, because the cognitive action of shape comparison requires the individual to pay attention to the components or the definitive properties of the observed shapes. Then this mental activity would follow with children distinguishing the similarities between different shapes in the same category. The perceived objects are not presented with their distinct properties since the individuation of properties requires attention to the relations among objects (Kelley, 1984). However, the ability to differentiate dimensions of similarity is the result of a developmental process, facilitated by examples and practice (Willingham, 2009). In the third step, children can categorize the similarities that they observed and form a concept that represents this category. In the last step, learners use this new concept that they formed in new situations. To recognize a shape accurately, children must be able to retrieve previously learned knowledge of shapes from memory and apply it to the current situation or a problem. Although the surface features of a problem change, the abstract operators used to solve the problem stay the same (Nokes & Ohlsson, 2001).

Building upon the cognitive and constructivist foundations established by Piaget, Dienes, and Mitchelmore and White, this study adopts a theoretical lens that bridges the cognitive process of abstraction with the perceptual and representational nature of geometric understanding. While Piaget and Dienes describe how abstraction develops, through empirical and reflective operations and through progressive engagement with multiple representations, this study adopts a constructivist lens centred on Fischbein's (1993) figural concept, extended by Vinner's concept image (Vinner, 1993), and Walcott et al.'s (2009) dynamic figural concept. This lens positions children's geometric thinking as a dynamic interplay between sensory images and formal properties, where understanding emerges through the construction, manipulation, and reorganization of mental and physical representations of shapes. It thus allows the study to focus not only on how children abstract properties from geometric experiences, but also on how their mental images, language, and embodied interactions (Maturana & Varela, 1987) evolve as part of that abstraction process. In this lens, children's geometric understanding is treated as sense-making-in-action, where the mental images they invoke and the properties they articulate co-evolve through manipulation, comparison, and talk; thus, the figural (sensory/imagistic) and the conceptual. With this orientation established, the following section formalizes the theoretical framework that integrates Fischbein (as the core lens) with Vinner's concept image and a constructivist reading of Walcott et al.'s dynamicity to guide the analysis of children's shape recognition and abstraction.

Theoretical Framework

The theoretical framework for this research is derived from an adaptation of Vinner's and Fischbein's perspectives. This study is based on the idea that the understanding of concepts in mathematics, and particularly geometry, involves an invocation of a set of mental images along with a corresponding set of properties for that class of objects. However, the framework is grounded in the understanding of the children's sense-making strategies. Concerned with the child's view of his or her mathematical world, my research lens encompasses the child's process of mathematical understanding through his or her sense-making process. I base my research on Fischbein's notion of figural concept through a constructivist philosophy (Walcott et al., 2009) that places the emphasis on students' actions on mathematical objects. In contrast to Fischbein's belief that the formal definition binds the figural concept, I consider a dynamic figural concept described by Walcott et al. (2009). It means that the child's dynamic figural concept develops within the natural development of the child's sensemaking strategies as he or she acquires mathematical experiences. The child modifies his or her visual prototype and/or verbal definition for a class of shapes as his or her perspective on the mathematical world evolves. These modifications are a direct consequence of the child's process of constructing meaning. When a child constructs new knowledge, he or she recalls an existing knowledge structure and refines or revises it to incorporate the newly acquired knowledge. Thus, the dynamic figural concept is comprised of the visual, verbal, written, symbolic, or formal properties of shape that an individual values (Mohr et al., 2007; Walcott et al., 2009). In my research I refer to the dynamic figural concept as the construct of images and definitions, along with representations other than visual or verbal, which a child holds for a particular class of shapes. For example, when an individual considers a rectangle as a dynamic figural concept, they will recall a collection of external representations, such as written illustrations, and internal representations, such as mental images, that he or she associates with rectangle that can be mentally manipulated and bounded by an explicit or implicit definition (Walcott et al., 2009).

While Vinner and Hershkowitz's (1980) concept image links the associated mental images and properties of a mathematical concept, Fischbein's (1993) figural concept is unique to geometric concepts. A figural concept binds a set of sensory mental images of a geometric figure, unique to an individual, to the formal definition of the figure. Fischbein's figural concept takes into account the mental manipulation of the internal object. He argues that a geometric figure is both characterized by a formal definition and a sensory image. This image, or "picture in the mind," possesses "spatial properties like extension, shape, location, magnitude" (Fischbein & Nachlieli, 1998, p. 1193). The formal definition of the figure and the sensory image combine to form the figural concept. Thus, when one considers a geometric figure, not only the formal definition is taken into account, but also the sensory image that the individual associates with that figure. The figural concept may be sparked verbally or visually. For instance, the sight of a parallelogram on the page sparks a sensory image of parallelogram that can be manipulated mentally and that is bounded by the formal definition of parallelogram. It is important to mention here that the figural concept is unique to geometric concepts and to everyone. In other words, one person's figural concept of rectangle may differ from another's depending upon the mental image(s) that he or she associates with it (Walcott et al., 2009).

In forming the theoretical framework for this study, I considered the assumption that the visual properties and physical attributes of shape work together in concept formation. Vinner and Hershkowitz (1980) suggested that understanding of mathematical concepts includes not only the understanding of the formal properties of the mathematical object, but also the variety of mental images associated with the concept; the concept definition, or "verbal definition that accurately explains the concept" (p. 177), often remains dormant in the cognitive process as children rely on the invocation of the concept image, constructed from the concept definition and common prototypes within the concept class. Vinner and Dreyfus (1989) refer to the concept image as "the set of all the mental pictures associated in the student's mind with the concept name, together with all the properties characterizing them" (p. 356). The mental pictures may include any type of representation, such as a picture, diagram, symbol, or graph (Walcott et al., 2009).

Fischbein's (1993) theory of figural concepts argues that a geometrical figure has intrinsically conceptual properties; however, it cannot be considered as a concept solely. Since it includes the mental representation of space property, it is also considered as an image. Therefore, according to Fischbein, all geometrical figures can be considered as mental constructs which includes conceptual and figural properties. The geometrical reasoning occurs through the interaction between the figural and the conceptual (Jones, 1998; Sinclair et al., 2016). According to Fischbein, the figural concept is the meaning that is inferred through the manipulation of words in a verbal activity. The mental activity of assigning words to the objects are the external, material representatives of meaning.

According to figural concept theory, a figure consists of spatial and sensory information structured by figural laws such as closure, proximity, and boundedness. A concept, however, is treated as an abstraction that is defined by its axiomatic properties. The geometry instruction must provide activities that help the individual to understand that a geometric figure is a figure and a concept simultaneously. It may be difficult to conceptualize a geometric figure if the figural aspects are not in line with the conceptual aspects of the figure, which may result in developing prototypical figural concepts. For example, the students may not recognize that a rhombus is a parallelogram even though they know the conceptual properties of a parallelogram (Sharma, 2019a, 2019b). The research suggests that the problems in children's shape recognition and categorization are associated with developing prototypical figural concepts that are static and isolated (Hallowell et al., 2015; Walcott et al., 2009). Although the children have the cognitive capacity to develop flexible conceptual geometric figures or transform their static prototypical concepts into flexible conceptual figures, there are certain problems affecting their shape recognition and categorization. While these theoretical models explain how abstraction operates cognitively, understanding how these processes manifest in children's actual geometry learning requires examining the persistent challenges they face in recognizing and categorizing shapes.

Problems Associated with Children's Shape Recognition and Categorization

Children as young as 12-months-old can learn categorizing based on a common feature and, form associations between elements in these categories (Gomez & Lakusta, 2004). Yet, abstraction is more than just categorization; it requires both compression and generalization. Compression occurs by zooming in on a complicated phenomena while focusing on essential and relevant features of interest, conceiving of these features as a whole, and making them available as an entity to think about (Gray & Tall, 2007). Generalization occurs by zooming out from a specific situation or object while recognizing that the features relevant to this specific situation or object may be ascribed to other situations, or objects.

Children’s capacity for visualization is crucial to comprehending geometric shapes. Children encounter a wide range of geometric shapes related to one another by shared characteristics. Visualizing the similarities and differences between the subsets of the shape categories improves student abstract thinking (Mohr et al., 2007; Walcott et al., 2009). Research shows that children tend to form and employ visual prototypes that represent classes of shapes as examples or as non-examples to identify and classify geometric shapes (Boswell & Green, 1982; Clements et al., 1999; Sinclair & Moss, 2012; Walcott et al., 2009; Yeşil Dagli & Halat, 2016).

Prototype is the term used to describe a conceptual example that is typical of its class. The prototype is essentially a mental image that is an outstanding example in terms of representation of the category in issue (Walcott et al., 2009). If we consider the categories or shapes in a radial structure, then the prototypes would be in the centre. Children base their conceptions on prototypes, which are the typical and most frequently utilized examples of the concepts (Lipovec, 2009). Prototypes form the core of an idea, and any other figures are categorized into the concept based on how much they resemble or relate to the prototypes. According to the literature (Aslan & Arnas, 2007a; 2007b; Clements, 2000; Hannibal & Clements, 2010), a typical triangle is perceived to be an equilateral triangle, whose lower edge is parallel to the ground on which it is placed. A typical square or a rectangle are perceived to have their lower edge parallel to the ground or page on which they are placed. Examples of triangles that differ in size, location, kurtosis (the ratio of height to the base) and skewness (the distance from the apex to the centre) are called atypical examples of the triangle (Aslan & Arnas, 2007b). The typical example is the central element of the triangle concept; all other varieties of a class of shape are related to it. Children's shape prototypes are formed based on the most frequent examples they encounter which represent a shape category (Arnas & Aslan, 2010; Hannibal & Clements, 2010; Lipovec, 2009). This means that if children are shown equilateral triangle more frequently than the other subsets of triangle as the typical example of a triangle, then their prototype would be an equilateral triangle. I begin by examining the first structural driver of these difficulties, limited exposure to atypical and non-example forms, which constrains the breadth of children's concept images and narrows their prototypes.

Lack of Exposure to Atypical Shapes

Aslan and Arnas (2007a) analyzed the content of 93 journals, 50 activity books, and 10 CDs for pre-school geometry education. Each shape in these resources was analyzed based on its intended use, whether it is there to help learning geometric shapes or learning basic mathematical skills such as classification, ordering, and matching, and whether it is a typical shape or not. The findings of the study show that the most frequently used shape was triangle, followed by square, circle, and rectangle, respectively. Also, 62.2% of the triangles were typical examples and 37.8% were non-typical examples. In addition, 64.1% of the rectangles, 70.1% of the squares, and 86.7% of the circles were typical examples. When the distribution of atypical examples is examined, it was seen that 14.5% of the triangles were represented in different sizes, 14.5% of the rectangle in different kurtosis, 21.6% of the square, and 13.3% of the circle were in different sizes. This means that the atypical examples that are shown to children mostly include different sized shapes with a slight emphasis on the kurtosis. This study emphasizes the importance of introducing children with typical, atypical and non-examples of shapes.

Children may pay attention to the shape's formal characteristics in addition to their informal or imposed characteristics, however, usually informal attributes applied to visual prototypes come before formal properties when children identify shapes (e.g., Burger & Shaughnessy, 1986; Clements et al., 1999; Halat & Yeşil Dağlı, 2016). Children implicitly determine whether a rotated or smaller-sized square is a square based on how much it resembles the central member of the square class, or family resemblance. This means that children compare shapes using prototypical judgement to determine class membership even when there is a verbal definition. In most cases, the formal definition of the shape and their visual prototype usually have no connection (Walcott et al., 2009). In studies with children aged 3–6, a huge percentage of children failed to classify non-prototypical triangles into the triangle category and rectangles into the rectangle category (Hannibal & Clements, 2010; Yeşil Dağlı & Halat, 2016). For example, 5-year-olds categorized some of the different triangles shown to them as half-triangles and combined some of them to resemble the isosceles triangle because they fit the triangle prototype in their minds. The most common triangle that they recognized was the equilateral triangle (Charlesworth, 2005). Even older children from 5th through 8th graders had trouble recognizing a twisted equilateral triangle with a base that wasn't parallel to the bottom of the page (Hershkowitz, 1989). However, these prototypes can be manipulated in ways that transcend shape classifications.

In other studies (Clements, 2000; Lipovec, 2009) triangles with an obtuse angle (>90°) were poorly recognized by children. Obtuse-angle characteristic appears to be more significant than baseline, non-symmetry, or aspect ratio in the identification of a shape. In terms of quadrilaterals, one trapezoid that has a right angle was more frequently recognized than the other trapezoid without a right angle which means that one of the square's most distinguishing characteristics, the model for quadrilaterals, is its right angle.

Fuson and Murray (1978; as cited in Sarama & DiBiase, 2004) reported that when children were 3 years old, more than 60% were able to recognize circles, squares and triangles. More recently, according to a study by Klein et al. (2004), 5-year-old children are 85% accurate in recognizing shapes, circle recognition, 80% in triangle recognition, 78% in square recognition, and 44% in rectangle recognition. About a quarter of 5-year-olds who were included in another study in Australia managed to recognize common shapes, including a series of triangles, before gaining any school experience (Sarama & Clements, 2009).

There are two types of prototypes: inflexible visual prototypes and flexible visual prototypes. Transforming inflexible prototypes to flexible prototypes is key to abstraction (Mohr et al., 2007; Walcott et al., 2009). Inflexible visual prototypes are the ones that are static and rigid in children's mind and are difficult to change. Children who formed inflexible visual prototypes are resistant to make changes on the shape. However, when learners can add motion to the static prototype they have, their prototype becomes a flexible one. Prototypes often become more flexible as learners gain a conceptual understanding (Lehrer et al., 1998). For instance, a parallelogram could be defined as a rectangle or an oval depending on children's implicit or explicit definitions about the shape (Archavsky & Goldenberg, 2005). In a survey of first-graders, 40% of the students suggested pushing on the corners to transform the rectangle into a parallelogram. Moreover, they rationalized the parallelogram as a "stretched rectangle" when asked to compare a rectangle and a parallelogram (Lehrer et al., 1998, p. 142). Being exposed to a wide array of shape classes with their subsets and conversating about their attributes, the similarities and the differences between these attributes as early as the preschool years, help children to stretch their prototypes to a more flexible one (Clements et al., 1999; Hallowell et al., 2015).

The features of the atypical examples of shapes that must be included in geometry education can be categorized in three groups:

- *Position* (orientation) is the stance of a figure in a place. When children see a geometric shape in a position different from its typical position, they may have difficulty identifying and classifying that shape accurately. Most of the preschool children think that the triangle should have a feature such as the vertex facing upwards. They may not include a triangle that is slightly tilted in the triangle category (Clements et al., 2004; Clements et al., 1999).
- *Aspect ratio* is the ratio of the height to the base in triangles and quadrilaterals. Children may have difficulty recognizing shapes when the height of triangles and quadrilaterals differs from their typical prototype (Clements et al., 2004; Clements et al., 1999; Kesicioğlu et al., 2011).
- *Skewness* is the distance of the vertex opposite to the base from the centre or the absence of symmetry; in other words, the distance of the vertex from the centre of the shape. Preschool children tend to exclude the triangles when the top point (vertices) is shifted to the right or left from the category (Clements et al., 1999; Kesicioğlu et al., 2011).

It is crucial for children to be introduced to a wide array of geometric shape classes (e.g., triangles, quadrilaterals, circles, and other polygons, prisms, cubes, spheres, and cylinders, etc.) with typical, atypical, and non-examples and their subsets (e.g., obtuse angled triangle, right angled triangle, isosceles triangle, pentagon, hexagon, heptagon, octagon, etc.) in order to further develop their concept images (Vinner & Dreyfus, 1989) and to form flexible prototypes (Walcott, et al., 2009). Beyond exposure, children also need opportunities to talk about and compare what they see; thus, we next consider how limited discourse about typical, atypical, and non-examples impedes definition-focused reasoning.

Lack of Discussion About Typical, Atypical, and Non-examples of Shapes

Children's early shape conceptions are not focused on the characteristics that define a specific class of shape, despite the fact that children are capable of categorizing shapes before learning their names (Quinn, 1987). Children may rapidly pick up on labelling shapes they come across; however, their shape categories are very primitive at the age of 2–3. Research suggests that children's initial referents for shape names are the most prevalent versions of those shapes such as an equilateral triangle on its base (Cross et al., 2009; Satlow & Newcombe, 1998). They do not consider the definitive properties that constitute the forms and do not apply shape labels to atypical variations of the shape class (e.g., that triangles have three sides and three angles) (Clements & Battista, 1992; Clements et al., 1999; Fisher et al., 2013; Satlow & Newcombe, 1998; Verdine et al., 2016).

According to the same research, children gradually expand their labelling of prototypical shapes to more and more atypical examples of shapes which leads to development of children's definition-focused conceptions (Verdine et al., 2016). To achieve definition-focused conceptual thinking, children must extend their labelling to atypical shapes. Though there is substantial disagreement regarding the extent to which language drives concept formation or vice versa, one hypothesis is that simply hearing shape names and seeing several examples might help children start to pick out the pertinent information among a range of properties, shapes, or objects (Gentner & Goldin-Meadow, 2003).

To help children extend their shape labels beyond visual perception-based matching to a limited prototype, discussion about geometric shapes of various sorts, their labels, and comparisons between shapes and its components in the same and distinct categories is probably necessary (e.g., Graham et al., 2010; Namy & Gentner, 2002). Sarama and Clements (2004) claim that early education teachers fall short of expanding children's shape knowledge during the short time that they spend discussing geometry. Rather than encouraging debate and discussion to promote understanding of the features of various shapes, they frequently ask for shape identification (e.g., What is this? What do we call this shape?) and then automatically confirm a right response.

The variations of shapes and the subsets of shape classes in educational materials are considered to be an important factor in refining children's conceptual understanding about the shapes. When there are varied subsets of a shape category, teachers use labels and statements of inclusion (e.g., "A rhombus is a quadrilateral") (Shipley et al., 1983) which prompt children to form wider categories of shapes (Liu et al., 2001; Shipley et al., 1983). Thus, if materials for young children included many varied shapes, parents and teachers might use language that highlights shape similarities and differences (e.g., "These are both triangles because they have three sides"). When teachers use everyday objects to teach shape names, they should do more than simply pointing out shapes (e.g., "look, a circle"). They should invoke the familiar object's name (e.g., "look that clock is also a circle) to highlight relevance and associations.

One study (Wiles & Anderson, 2019) investigated the impacts of discursive framework on geometric learning of 4th graders and instructional experiences of their teachers. During the study, teacher-researchers showed classes of geometric shapes to the students and asked the class what made a shape parallelogram. Most of the students answered that it had four sides. Even though the students' answers were not very reflective of all the aspects of parallelogram, the discursive approach where the learners and the teachers discuss the definitive aspects of shapes uncover learner's conceptual understandings about the shapes. Verdine et al. (2016) suggest that guided play with geometric objects where the teacher and the children construct geometry knowledge together by conversating, comparing geometric shapes in terms of their similarities and differences, and using distinctive language to label each property or shape can potentially promote children's conceptual understanding of the shapes.

It is important for children to distinguish between the names of two-dimensional shapes and the names of three-dimensional shapes and label them correctly. However, in a study by Downton and colleagues (2019) in which primary school children were shown different geometric shapes and asked to identify them, they misnamed some of the shapes. Although children understood and remembered the properties of the object discussed, it was found that they named the solid materials differently than expected; for example, they used names like "square" for the cube (Downton et al., 2019). Such problems indicate the importance of using correct definitions when teaching geometry to children (Dauksas & White, 2014).

In a U.K. study by Shorrocks et al. (1991; as cited in Sarama & DiBiase, 2004), 7-year-olds were given a series of shapes in random order and asked to identify them by matching them. More than 90% easily mapped shapes, but the accuracy of these matches is 97.4% for the circle, 96.4% for the square, 92.8% for the rhombus, 78.1% for the rectangle, 55.3% for the hexagon, and 31.1% for pentagon. At the same time, it was observed that the children had a confident attitude towards drawing the shapes that they had difficulty naming. When asked to explain the features of the figures, 92% of the children were able to explain the characteristics of the square and 80% of the pentagon. These data show that children can easily distinguish shapes, but are introduced to a limited number of shapes. Although these children's performances are lower for foreign shapes, they can define the features of some shapes (Orton, 1996). This indicates that the education given to these children was insufficient in terms of geometric information. Such research emphasizes the importance of introducing these shapes with their correct names, as well as the variety of shapes when teaching children geometry. It is important to ask children to describe shapes using the vocabulary they currently have. Repeating description activities throughout geometry instruction over time can help foster abstract thinking. Describing shapes encourages children to actively use mathematical vocabulary and connect it with their conceptual understanding. Teachers can see the vocabulary that the children are using and slowly introduce more abstract and appropriate terminology within their activities and conversations with children. Even when talk is present, progress depends on directing attention to the properties that actually define classes, so the following subsection addresses children's limited experience attending to identifying features across varied instances.

Lack of Experience Attending to the Identifying Features of Typical, Atypical, and Non-examples of Shapes

According to Fischbein (1993), the ideal representation of a class of objects based on their shared characteristics is a part of the conventional notion of a concept. From this perspective, geometrical concepts refer to common properties of a class of geometrical shapes or objects that can be mentally visualized or perceived when interacting concrete representatives. For example, some figural characteristics, such as the surface shapes of an object or the angles that control how the surfaces are connected, may indicate that an object belongs to a particular class such as prisms, cubes or spheres. As children assemble and disassemble three- or two-dimensional objects with manipulatives such as rods, beads, and/or plane tablets, they use the same materials to assemble different shapes or objects. The same manipulatives (e.g., beads, rods, planes) that they use to construct different shapes (e.g., square, triangular prism, rectangle) are the common properties or common components between those figural concepts. Assembling and disassembling two- and three-dimensional shapes which were perceived as a one-piece figure have the potential to facilitate children's understanding of the relationship between the shapes. Walcott et al. (2019) found that elementary school children who identified the shapes as different generally focused on the classification of the shapes or on attributes of the shapes; however, the children who focused on attributes of shapes decomposed the figure into what for them were critical features in the comparison. Many children describe triangles as shapes with "three dots and three edges" but half of these children do not know exactly what the concepts "edge" and "dot" are (Clements et al., 1999). This means that using activities such as composing, de-composing, and transforming two- and three-dimensional shapes have the potential to direct children to focus on the components of the shapes.

Early geometry education must assist children in understanding how the shape is composed, which relationships can be established within the shape and its components and how these work together when composing, de-composing, or re-composing a shape (Conceição & Rodrigues, 2022). Research (Kamii et al., 2004; Sarama & Clements, 2009) shows that block building has a positive impact on the development of understanding of shape and shape composition, thus also facilitating children's general reasoning about the geometric shapes. One study found that construction, composition, and de-composition activities with blocks help create the groundwork for children's arithmetic success in the elementary years (Wolfgang et al., 2001). Hallowell et al. (2015) investigated and explained Grade 1 students' difficulties in relating two-dimensional shapes with three-dimensional objects. The study underscored the importance of giving children adequate hands-on experiences and opportunities for dialogue during composition and decomposition activities. Previous learning experiences with shapes seem to influence students' performance in tasks involving de-composition and re-composition (Spiegel & Ginat, 2017). Sinclair and Bruce (2015) argue that there are many instances of three-dimensional shapes in early grades' classes, however there are not enough studies that investigates how the students learn about these shapes. Clements and Sarama (2020) propose two learning trajectories for the composition of two- and three-dimensional shapes. One trajectory focuses on relationships between components and composites, while the other emphasizes relationships among components, composites, and the whole. Beyond children's cognitive difficulties, the literature also highlights how teachers' limited geometric understanding and instructional practices contribute to these challenges in classroom settings.

Problems Associated with In-service and Pre-service Teacher's Knowledge and Experiences in Geometry Teaching

It is critical to understand the present state of teachers' geometry knowledge, how it improves, and how professional development might promote this knowledge development. Schulman (1986) distinguished the subject matter knowledge referring to key facts, concepts, principles, and explanatory frameworks of a discipline, as well as the rules of evidence used to guide inquiry in the field from pedagogical content knowledge which refers to understanding of how to present specific topics in the most effective and appropriate ways to the learners. Research (Clements, 2003; Fujita & Jones, 2006) shows that many teachers consider the overall figure of a shape when recognizing shapes, instead of paying attention to the definitive properties of geometric shapes. This indicates that if teachers themselves don't refer to the properties of the shapes as identificatory or definitive characteristics, they also would not likely teach these properties within their activities.

According to research (Jones, 2000; Jones et al., 2002), it would be beneficial to pay attention to math teachers' initial and ongoing education in terms of their background and understanding of geometry, as the effective teaching of geometry depends on teachers having a thorough understanding of geometry and knowing how to teach it effectively. For instance, many aspiring teachers are only able to recognize and categorize shapes at Level 1 of the geometric thinking model developed by van Hiele (1986; as cited in Teppo, 1991) (e.g., “that must be a rectangle because it looks like a door”); Level 2, the descriptive/analytic level, where people recognize and categorize shapes based on their characteristics, is often not taught to many people (Clements, 2003; Swafford et al., 1997). For instance, in one study (as cited in Sarama & Clements, 2009), prospective teachers in the U.K. were only capable of Level 2 reasoning. It is acknowledged that having a thorough understanding of mathematics topic or subject knowledge alone is not sufficient to teach mathematics effectively, as the pedagogical knowledge of mathematics is as crucial as the content knowledge itself (Jones et al., 2002). To address these interconnected issues, scholars have revisited historical yet pedagogically rich tools such as the Froebel Gifts, which offer structured, manipulative-based experiences aligned with principles of abstraction and geometry learning.

Froebel Gifts

Froebel Gifts introduce the mathematical concepts and operations with a specific alignment moving from three dimensional to two-dimensional, surface (plane) to line, line to point. Gift 1 consists of six balls; each are connected to a string and highlights the similarity and dimensions of similarity in shape and colour. As all balls have the same characteristics except colour, the attention of the child would be directed on different colours. Three- and 4-year-old children can understand the concepts of right-left, up-down, on-under, here-there, near-away, front-behind, using the ball. Moving, bouncing, jumping, rolling, bending, falling, pulling, and pushing are exemplified by the movements of the ball, giving the child an opportunity to learn spatial orientation (Eugene & Provenzo, 2009). Gift 2 consists of two 5 cm cubes, a cylinder 5 cm in diameter and height, and a wooden sphere 5 cm in diameter. The gift is designed to highlight that the cylinder has both rectangular and circular properties. However, the sphere is completely circular, and the cube is completely rectangular. This gift provides the opportunity to compare circular and rectangular shapes and their properties.

Gifts 3 to 6 consist of a variety of three-dimensional geometric shapes such as cylinders, spheres, cubes, oblong blocks, rectangular prisms, triangular prisms, narrow columns, and cuboids. Children can observe that these blocks with different lengths, width, or height constitute bigger geometric shapes through teacher's guided instructions. For instance, in the Gift 3, children can see that when eight cubes are put together, they constitute a bigger cube which is completely identical with the small cube in terms of its shape, however, different in terms of size. Gifts 4, 5, and 6 are given to children respectively after Gift 3, with the aim to facilitate the development of fractions in mathematics with various divisions, gradually. This alignment has the potential to lead the child to experience fractions with blocks varying in size and shapes. For example, with Gift 5, teachers initially initiate activities such as separating the cube into congruent thirds in different ways, then into ninths. The components of each gift between Gifts 3 and 6 (e.g., the cubes in Gift 3, rectangular prisms in Gift 4, the triangular prisms in Gift 5, and oblong blocks in Gift 6) constitute a cube. The cube that they constitute is identical in shape and size, however, different in number of its constituents. This experience, through the manipulation of materials, provides an opportunity for children to experience the concept of fraction by intentionally omitting the inconstant properties such as the specific shape or number of constituents and see the similarities among the operations. The first six gifts objectify the concept of a fraction by providing the children with the opportunity to count the blocks that constitute a bigger geometric shape such as a cube or a rectangular prism and compare the units with the whole of the shape in terms of properties like size, and shape. Gift 3 emphasizes the multiplication, size, and fractions by comparing the whole and parts. Gift 4 emphasizes that rectangular parallelepiped and cuboids are the same in terms of the number of edges, vertices, and faces, however, they are different in terms of edge length. The comparison opportunities that Gifts 5 and 6 provide is that cube, oblong prism, rectangular parallelepiped, and cuboids are more similar in terms of the number of edges, vertices, and faces than triangular prism (Reinhold et al., 2017). Froebel Gifts 3 to 6 represent the multiples of two, three, and four with parts of different sizes and shapes, allowing the child to grasp the concept of the unit in a relative and varied alignment of objects. This process represents the division process in a concrete analogy and is designed to lead to multiplication operations and fractions.

Gift 7 introduces not only typical examples of geometric shapes, but also atypical examples (Marenholtz-Bülow, 2007). The shapes of the parquet tablets in Gift 7 are obtained from the surfaces of the 3D objects in the first six gifts. With Gift 7, the child may perceive surfaces as part of solids. The flat tablets represent the concept of surface or plane of different 3D objects. Therefore, this gift is important for the transition from three-dimensional to flat surfaces (planes). The first six gifts allow the child to create mental images of real three-dimensional objects such as cube, cylinder, sphere, and prisms (Ahmetoğlu & Ildiz, 2018; Correia & Fisher, 2014; Palmer, 1912). Gift 7 allows the child to create the abstract concept of two-dimensional forms of three-dimensional objects. With the activities in Gift 7, children can compare different shapes in general and then compare similar shapes to find differences between them.

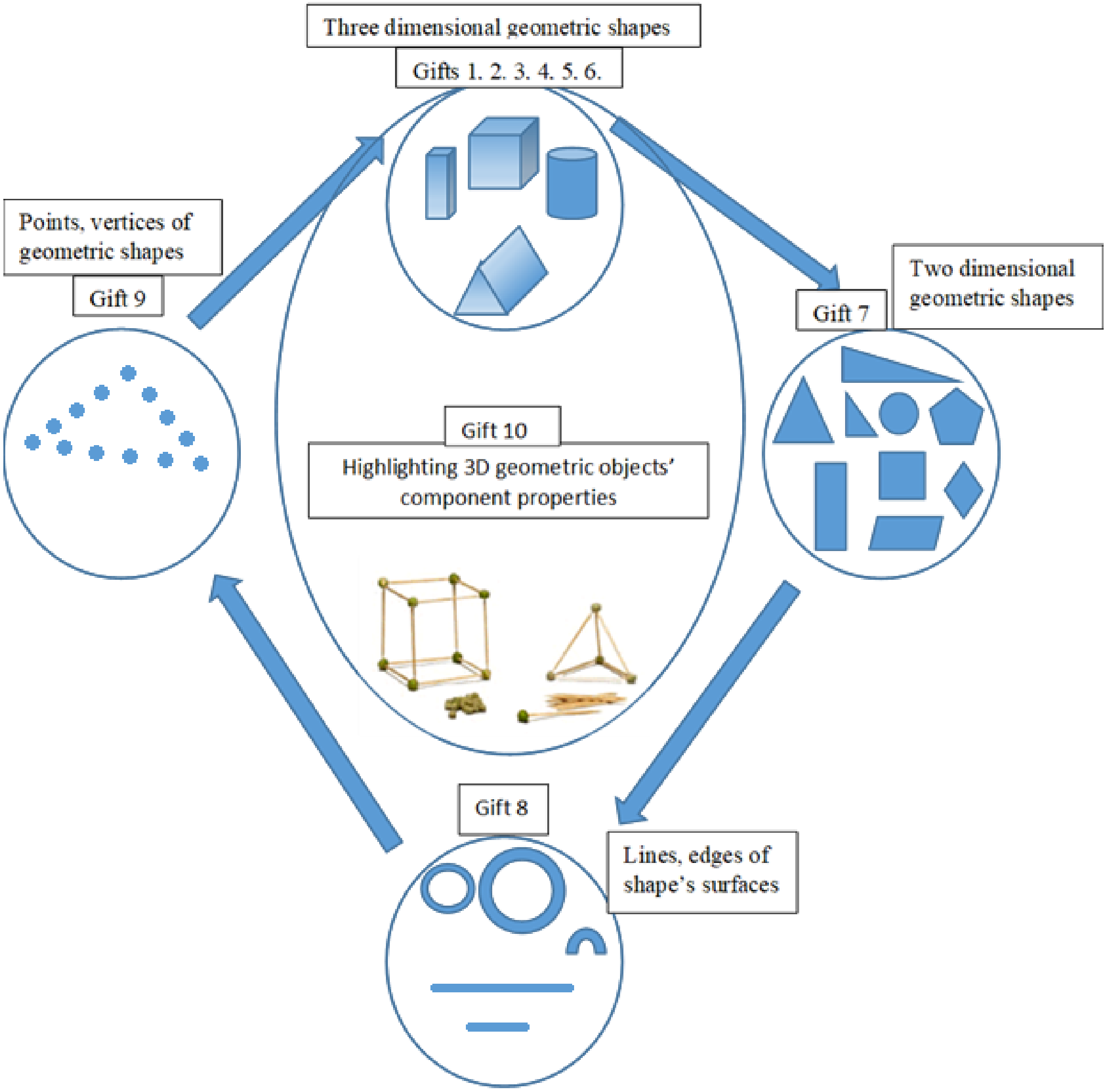
Gift 8 symbolizes the transition from the plane (surface) to the line concept, embodying the boundaries, edges, and lines of planes and three-dimensional objects. In this gift, the concept of length is particularly emphasized, and the most characteristic feature of the gift is that it deals with the concept of straight or curved lines (Correia & Fisher, 2014). The alignment of concepts in FG is designed to proceed to the concept of point with Gift 9. This gift introduces the point on its own as the most basic building block of previous gifts. In fact, although there is no length, depth, or width of the point, it is transformed into a tangible object for the child to understand it in a concrete way (Wilson, 1967). Gift 9 provides an opportunity to the child to comprehend the concept of straight line and that a line is the shortest distance between two points and changes direction at each point of a curved line. All planar or curvilinear geometric shapes or all angles can be represented with dots (Woodard, 1979). The gifts are designed to be gradual, from the blocks that represents 3D shapes, to the parquet tablets that represents planes, continuing with the rods that represents lines and beads that represent the point.

With the last gift, Gift 10, abstract concepts such as shape, form, area, line, and point can take a solid and compound form. For this experience, the use of all previous gifts must be internalized (Correia & Fisher, 2014). Gift 10 is fundamentally important in the abstraction processes in shape recognition, because it provides the opportunity to compose and de-compose three- or two-dimensional shapes by using rods as sides of the shape, beads as vertices, and tablets to cover the surfaces of the skeleton of the shape. In this process, children can possibly treat the three-dimensional shapes as geometric forms consisting of components such as sides, vertices, angles, and planes, instead of regarding them as overall visuals. Figure 2 displays how the Froebel Gifts are designed to introduce the geometric concepts with a specific alignment moving from three dimensional to two dimensional, from surface to line, and from line to point.

The richness of Froebel Gifts in terms of geometric diversity may carry the potential to enable children to meet not only typical examples of geometric shapes but also atypical examples. For example, the triangle diversity of Gift 7 has the potential to enrich the triangle category of children and allow them to make the definition of the triangle more accurately. When children were taught using guided play (Hirsh-Pasek et al., 2009) as opposed to didactic instruction, they were dramatically better in transferring their shape knowledge to atypical shapes not previously seen (Fisher et al., 2013).

Figure 2

The Order of Highlighted Geometric Concepts in Froebel Gifts



Research shows that when there are varied instances of a category (e.g., square, rectangle, rhombus, trapezoid, etc.), adults use the correct terminology (e.g., “A rhombus is a kind of parallelogram”) that helps children to develop wider categories (Shipley et al., 1983). Groups of objects bring out super ordinate vocabulary, and this process helps the children to classify objects as members of larger classes (Liu et al., 2001). This indicates the importance of variety in geometric shapes, more precisely the varied instances of each shape category, because if materials for young children included many varied shapes, parents and teachers might use language that highlights shape similarities and differences (e.g., “These are both triangles because they have three sides”). They might also be more likely to signal the existence of wider categories (e.g., “This weird one is a kind of triangle.”). Yet popular shape sorter toys and touchscreen apps typically contain only one, iconic version of each shape category (Verdine et al., 2016). Furthermore, designing shape materials requiring goal-directed adult involvement (e.g., Ferrara et al., 2011) might result in significantly more exposure to language about geometric forms at younger ages (Verdine et al., 2016).

In the research carried out by Arnas and Aslan (2005), it was observed that most of the children in early childhood defined the shapes and matched them with a visual prototype. At the same time, it was revealed that the shape they had the most difficulty in classifying was triangular (Arnas & Aslan, 2005, 2010). Young children tend to accept long parallelograms or right-angled trapezoids as rectangles. Thus, we can say that a child’s visual rectangular prototype is a nearly square shape in terms of its corners with two long parallel edges. However, because Froebel’s Gifts present a variety of geometric shapes, children can see rectangles and trapezoids side by side and learn to give each figure a distinct name through structured instruction and comparison tasks. This would allow for children to expand their geometric shape categories and the geometric vocabulary due to the name given to each different shape.

Materials that introduce the relationship between the names of shapes and the features of the shape to the children, and more importantly, allow children to discover these relationships spontaneously, and a training program compatible with these materials may be the solution to the geometrical deprivation found in the research. In the training program in which Froebel Gifts are used, the differences and similarity relationships between the various shapes, the connection between the name and the features of the shape, and the categorization of atypical geometric shapes are important.

In the training program using Froebel Gifts, the transition from three dimensions to two dimensions was supported with meaningful examples, and the names of the 3D shapes and the names of the 2D shapes were introduced by emphasizing the differences between them. There is a gradual transition from three dimensions to two dimensions in Froebel Gifts, especially from Gifts 6 and 7. In all gifts up to the Gift 7, the edges, surfaces, and corners of three-dimensional objects are counted and compared. Moving to the concept of surface with Gift 7, the edges of the surfaces are counted and the differences with the three-dimensional shapes are focused on. This process helps the child to visually understand the difference of two and three dimensions and embody the knowledge through the senses' manipulation and language articulation.

When children reach the age of 5 or 6, they can recognize basic geometric shapes. They can define the shape with their own sentences, compare the shapes, and classify them according to their specific features. They can also visualize and draw a certain shape. They have limited success in researching and defining differences between geometric shapes (Verdine et al., 2016). However, using these blocks provides children with concrete opportunities to make discoveries about two-dimensional and three-dimensional shapes. Children discover various shapes in spontaneous arrangements. Children also make many discoveries about shapes by constructing objects while playing creative games (Hirsch, 1996, as cited in Copley, 2000). The fact that children identify patterns in space or realize the relationship between the surfaces of three-dimensional shapes, the number of sides and corners leads to their understanding of the basics and functions of algebra. As children compare shapes, directions and positions in space, they develop new concepts and feel the need to learn the necessary points for these concepts to use in measurement processes. Classifying elements by shape or other geometric feature is a basic skill for data collection. The role of the teacher is to bridge the informal knowledge gained by the child's own experiences with the official mathematics content. This bridge usually means using the child's own language to associate the child's knowledge with formal terms and definitions (Clements, 2001).

Children first try to recognize geometric shapes and name them by looking at their appearance. Then they distinguish the features of the shape. For example, they notice that the circle has no corners, or the square has four edges. They define the features of the shape with their own unique expressions. For example, they indicate that the square consists of four straight lines. Finally, they establish relationships between the features of the shape. For example, they can explain that the square is included in the rectangular class because it has four edges. At this point, they are introduced with geometric terms and start using words like edge, corner, surface. In this process, it is necessary to provide children with opportunities to freely explore two- and three-dimensional shapes. Blocks, legos, and similar materials provide suitable opportunities for this discovery (Charlesworth, 2005; Copley, 2000).

When preschool children define geometric shapes, they define and classify them according to the main features they consider for each shape. Geometric shapes have some decisive and non-decisive features. The edge, number of corners, and edge length of a geometric shape are the defining features of that shape, because it is these features that distinguish a rectangle from a triangle or square. The position, skew, or kurtosis (height-to-base ratio) of the shape are non-determining features for geometric shapes, because even though a triangle is crooked, it is a triangle. It has been observed that preschool children perform the defining and classification processes of the shapes, while they consider the non-determining features of the shapes more often, but they also pay attention to their determining features (Arnas & Aslan, 2007b). The Froebel Gifts, however, are most effective when implemented within the broader pedagogical philosophy of the Froebel Approach, which integrates guided play and conversation to foster abstract and reflective thinking in young children.

Froebel Approach

As children's cognitive skills develop, they can see even part of an object, for example, one of the vertices in a triangle, and know that it is part of a whole object. When children have opportunities to explore two- and three-dimensional objects, they develop an ability to coordinate movement and alignment of those objects (e.g., putting a square tablet on a square hole). When these activities are guided by the teacher, children also learn the terminology used in spatial thinking and spatial activities (Ahmetoglu & Ilhan Ildiz, 2018; Tovey, 2016; 2019). During these activities, the spatial language that the teacher introduces to children (e.g., angles, lines, vertices, planes, long, high, side, angle, same, symmetrical, triangle, rectangular prism, area, etc.) by using Froebel Gifts provides them with essential tools to describe the shapes accurately and more abstractly. The teacher pays special attention to use the correct words for the shapes and the properties of shapes while using the related tangible objects (e.g., when talking about the pointy corners of a cube, she/he uses the term "vertice" and uses the beads in Gift 10 to represent the vertice). Starting from Gift 1, the teacher asks children three forms of questions:

- Forms of life: Questions about what the shape at hand (e.g., cube, triangular prism, sphere, etc.) resembles in life (e.g., a tree, a car, a glass, a tree trunk, etc.).
- Forms of knowledge: Questions about the properties of the shape (e.g., the size of angle, the number of sides, the length of sides, the number of vertices, the shape of planes, the number of surfaces, etc.).
- Forms of aesthetics: Questions about symmetry, beauty, and aesthetics (e.g., Are these two shapes symmetrical? Can you construct me a beautiful shape?) (Eugene & Provenzo, 2009).

During the conversations between the teacher and the children, the teacher emphasizes and reinforces the usage of the correct vocabulary about the shapes and the properties of shapes. If necessary, the teacher corrects the children and reinforces them to use the correct vocabulary in their explanations. This exercise is used as a metacognitive mediator for children to internalize the precise verbal definitions of geometric shapes. Vygotsky (1978) suggests two major types of mediation as the main mechanism of children's learning and development as metacognitive and cognitive. Metacognitive mediators are semiotic tools of self-regulation: self-planning, self-monitoring, self-checking, and self-evaluating (Bartolini & Mariotti, 2008; Karpov & Haywood, 1998). As teachers consistently use the correct vocabulary when talking about the shapes and corrects children and reinforces them to use the correct vocabulary, they also regulate the children's inner speech, because the words and the speech that the teacher uses are internalized by children. The egocentric inner speech that the child uses is the first steps of self-regulation.

Froebel Gifts can also be used as a cognitive mediator to facilitate children's acquisition of cognitive tools that are necessary for solving spatial problems. Cognitive mediation refers to the understanding of scientific concepts that represent the fundamental aspects of some class of phenomena (Karpov & Haywood, 1998). Preschoolers tend to use spontaneous concepts instead of scientific concepts if they are not introduced to these scientific concepts through systematic instruction. The spontaneous concepts are formed as a result of generalization and internalization of everyday personal experience in the absence of systematic instruction. According to Vygotsky (1978), the acquisition of scientific concepts should arise from their presentation to children in the form of precise verbal definitions (Duncan & Tarulli, 2003; Karpov & Haywood, 1998). When teachers use Froebel Gifts, they explain the meaning of the geometric vocabulary and help children make sense of these verbal definitions (e.g., the word "triangle" is a combination of the words "three" and "angle" which means that the shape has three angles). Building upon these Froebelian principles of guided exploration, linguistic mediation, and progressive concept formation, I developed the Froebel Gifts Early Geometry Program (FG-EGP) to translate these theoretical foundations into a structured sequence of geometry learning experiences that move systematically from concrete manipulation to abstract reasoning.

Froebel Gifts Early Geometry Program

I designed the FG-EGP as a structured, play-based geometry intervention that follows a progressive sequence from concrete to abstract representation. The program begins with children's exploration of three-dimensional shapes and gradually transitions toward two-dimensional figures, lines, and finally points. This ordered progression mirrors Froebel's pedagogical principle of unfolding complexity, helping children perceive how geometric entities are related and how higher-order abstractions emerge from tangible experience. At each level, activities are designed to draw children's attention to structural similarities and differences, facilitating the recognition of defining geometric properties rather than superficial visual features.

Central to the FG-EGP is the use of guided comparison as a vehicle for abstraction. Children engage in side-by-side explorations of shapes, comparing three-dimensional solids with their corresponding two-dimensional surfaces, differentiating among two-dimensional shapes that share or vary in certain properties, and even relating lines to points. These comparisons are intended to help children notice invariant properties across perceptually different examples. For instance, when contrasting a cube with its square face or a triangle with another triangle of different orientation, children learn to focus on the defining features such as number of sides, vertices, and angles. The teacher scaffolds this process through purposeful questioning that encourages reflection and articulation ("What makes these shapes the same?" or "How is this one different from that one?"), gradually guiding children toward property-based reasoning.

Another key feature of the program is its emphasis on composition and de-composition as active learning processes. Children are encouraged to construct shapes from smaller components and to decompose complex figures into simpler parts, using the Froebel Gifts manipulatives as tangible tools for reasoning. These hands-on construction tasks promote understanding of geometric relationships, such as how surfaces are bounded by lines, or how lines connect at points, while also supporting the development of spatial reasoning, symmetry, and part–whole understanding. Through repeated cycles of building and breaking down shapes, children internalize geometric concepts as dynamic, interconnected systems rather than fixed images.

Although the FG-EGP emphasizes guided instruction, each session begins with a period of free play to allow children to explore and familiarize themselves with the materials. This unstructured time encourages curiosity and sensory engagement, laying the groundwork for more focused, teacher-guided exploration. The guided component that follows transforms children's spontaneous discoveries into opportunities for conceptual understanding, using dialogue, reflection, and modeling to connect physical manipulation with formal geometric vocabulary. This combination of free and guided play enables both creative exploration and systematic concept formation, providing a balanced approach that supports early abstraction in shape recognition.

In sum, the FG-EGP integrates Froebel's sensory–constructive philosophy with contemporary understandings of guided play and cognitive development. By engaging children in progressive comparisons, active construction and deconstruction, and reflective dialogue, the program fosters a gradual internalization of geometric relationships, from surface perception to abstract reasoning. This design allows children to move fluidly between perception and conception, forming flexible, property-based understandings of shape that underpin the development of early geometric abstraction.

In summary, the literature reviewed in this chapter establishes that young children's abstraction in geometry is shaped by the interplay of perceptual prototypes, language, and repeated opportunities to compare, compose, and decompose shapes within meaningful social interaction. It also identifies Froebel Gifts and the Froebel Approach as a pedagogically coherent response to common barriers in shape recognition, while positioning Fischbein's figural concept and Vinner's concept image as an analytic lens for examining children's evolving geometric sense-making. Building on these theoretical and empirical foundations, Chapter 3 outlines the methodology of the study, including the research design, participants and setting, the structure of the FG-EGP intervention, data sources and procedures, and the analytic strategies used to investigate children's abstraction processes in shape recognition.

CHAPTER THREE: METHODOLOGY

In this chapter, I describe how I investigated two research questions:

1. How did differences in intentionally designed geometry activities, vocabulary, and manipulatives used by teachers impact 3- to 4-year-old children's abstract thinking skills in shape recognition?
2. What kind of activities, vocabulary, and manipulatives did teachers use in their geometry activities before, during, and after the intervention?

The first question served as the primary focus of the research, while the second question was designed to collect supplementary data that could provide deeper insight into the main question. In the following sections, I explain the logic of this design and the rationale for choosing an embedded mixed-methods model. I then describe how I implemented the study in four daycare centres with nine classrooms, involving 93 children and nine teachers. I outline how I conducted the intervention, including the duration differences across classrooms (4 or 8 weeks), and how I organized each phase from participant recruitment to pre-testing, training, implementation, post-testing, and interviews.

I also explain the tools and procedures I used for data collection and analysis. I administered the Geometric Shape Recognition Test (GSRT) individually to measure both recognition accuracy and property-based reasoning, conducted structured classroom observations to document instructional practices, and carried out teacher interviews to gain insight into pedagogical intentions and reflections. Finally, I describe how I analyzed the quantitative and qualitative data and what steps I took to ensure validity, reliability, and trustworthiness. By structuring the chapter in this way, I aimed to present a transparent account of how I designed, implemented, and analyzed the study to examine the effects of the FG-EGP on young children's abstraction in geometric shape recognition. To address these questions rigorously, I first justify the overall methodological stance and explain why an embedded mixed-methods design is most suitable.

Choosing a Mixed-Methods Research Model: Embedded Experimental Design

Given the nature of the research problems and the intent to investigate both outcomes and processes, a sequential embedded mixed methods design was selected. As outlined by Creswell (2006), mixed methods research may follow various models, such as triangulation, embedded, explanatory, or exploratory. In this study, the embedded design was most appropriate because the qualitative data served to support and deepen the interpretation of the quantitative results.

An embedded design was chosen because of its capacity to examine intervention effects while also providing process-level insights. Following Creswell and Plano Clark (2018) and Fetters et al. (2013), an embedded design nests one method inside another to address different questions within a single study, typically giving priority to one strand (the "core"), while a secondary strand supplies supportive or explanatory information collected before, during, and/or after the core intervention. In this study, the core quantitative strand tested the impact of the FG-EGP on shape recognition scores and explanation types (property-based vs. visual) using pre/post measures and group comparisons, while the nested qualitative strand, two rounds of observations (pre- and during-intervention) and post-interviews, documented process, implementation, and mechanisms (e.g., vocabulary emphasis, use of Froebel Gifts, comparison/composition/de-composition activities, teacher scaffolding).

A triangulation design was ruled out due to the logistical limitations of conducting equal-priority methods simultaneously as a sole researcher. An exploratory design (QUAL → quan) was also not appropriate since the study did not aim to develop a new instrument; validated tools such as the GSRT (Aslan & Arnas, 2007b) were already available. The embedded experimental model, therefore, allowed for an integrated understanding of how structured instructional factors, such as vocabulary use, manipulative handling, and teacher scaffolding, contributed to changes in children's geometric abstraction. Having established the rationale for an embedded model, I next detail the specific research design and how the strands were organized across sites, groups, and time.

The Research Design

Mixed methods research supports investigations that transcend methodological boundaries, enabling the researcher to focus on research questions rather than philosophical allegiances (Schoonenboom & Johnson, 2017). By collecting and analyzing textual, verbal, and numerical data, this study provided a holistic picture of how children develop abstract thinking in geometric shape recognition. The strengths of each methodology compensated for the limitations of the other, allowing for a triangulated understanding of how geometry instruction influenced learning.

Following ethics approval from the Brock University Research Ethics Board (REB), data were collected from 93 children and nine teachers across nine classrooms in four licensed daycare centres located in the Niagara Region of Ontario, Canada. All participant names reported in this study are pseudonyms to protect confidentiality.

Quantitative data were gathered through the GSRT, which was administered to children in both experimental and control classrooms before and after the intervention, which assessed changes in children’s recognition of typical and atypical geometric shapes, and their coded verbal responses measured abstraction. The quantitative component examined growth in abstract thinking skills across four groups, two experimental and two control, by analyzing differences in

GSRT scores over time and across conditions. The independent variables were group (experimental or control) and duration (4-week or 8-week intervention), and the dependent variables were children's shape recognition scores and their verbal responses coded for abstraction (e.g., property-based descriptors). This component employed a quasi-experimental pretest–posttest control group design. The experimental classrooms implemented the FG-EGP, developed by me to promote geometric abstraction, while the control classrooms continued with their regular geometry instruction and materials.

The qualitative component aimed to explain how and why this growth occurred by analyzing classroom practices, vocabulary use, and manipulative-based activities through observations conducted before and during the intervention, and teacher interviews conducted after the intervention to provide insights into children's engagement, vocabulary use, and teacher strategies for supporting abstraction. These data were thematically analyzed to understand the pedagogical mechanisms behind children's abstraction processes and to identify patterns of practice within and across groups. With the design in place, the following section describes the participants and classroom contexts to clarify who received which condition.

Participants

After obtaining research ethics approval, I began the recruitment process by sending email invitations to all 180 licensed daycare centres in the Niagara Region, using the list of day cares.[1] Several centres responded to the invitation, and I continued the process with those that demonstrated both a strong willingness to participate and appropriateness for the research context. Appropriateness was assessed based on factors such as the demographic similarity of the communities served, availability of classrooms with 3- to 4-year-old children, and the logistical

---

[1] See: https://www.parentdirectniagara.ca/cat/licensed-child-care-centres

feasibility of completing all research phases. Ultimately, I selected four daycare centres to participate in the study. I secured consent from participating teachers and parents. Once consent was obtained from administrators, teachers, and parents, and assent from children, classrooms were assigned to either the experimental or control condition with efforts to minimize confounding variables. Children's verbal assent was obtained at the beginning of testing. Final selection of classrooms for participation was based on administrator and teacher input regarding the learning environment and community culture. Four daycare centres were chosen, resulting in a total of nine classrooms: four experimental (one 4-week and three 8-week) and five control (four 4-week and one 8-week).

The research involved a total of 93 children and nine teachers across nine classrooms. Two centres hosted both experimental and control classrooms, one was exclusively experimental, and one served entirely as a control site. While five classrooms followed an 8-week intervention schedule, the remaining four classrooms had a shortened 4-week implementation due to holiday closures, allowing time duration to be tested as a moderating factor in the analysis. The research process unfolded over 16 weeks and was structured into three key phases: (a) pre-intervention observations and pretesting, (b) intervention and ongoing observations, and (c) post-testing and teacher interviews.

Table 1 shows the distribution of participants based on their age, gender (female and male), and group (control and experimental 8 weeks; control and experimental 4 weeks). The majority of participants (64.5%) were 3 years 9 months old to 4 years 3 months old, with fewer participants 3 years 3 months old to 3 years 8 months old and very few 4 years 4 months old to 4 years 10 months old. This distribution reflects the age range of the study group, with most participants likely closer to 3 years old. The smaller group (older participants) could affect the generalizability of findings for that subgroup.

Table 1

Descriptive Characteristics of Child Participants by Age, Gender, and Group Assignment

Factors	N	%
Age		
3 years 3 months old to 3 years 8 months old	25	26.9%
3 years 9 months old to 4 years 3 months old	60	64.5%
4 years 4 months old to 4 years 10 months old	8	8.6%
Gender		
Female	48	51.6%
Male	45	48.4%
Group		
Control group (8 weeks)	28	30.1%
Experimental group (8 weeks)	29	31.2%
Control group (4 weeks)	21	22.6%
Experimental group (4 weeks)	15	16.1%

The study group is almost evenly distributed by gender, 48 females (51.6%) and 45 males (48.4%). This situation ensures that any findings are unlikely to be biased due to gender imbalance. This balance allows for comparisons or subgroup analyses by gender. Participants are distributed across experimental and control groups and based on the duration of the intervention (4 or 8 weeks). The groups are relatively balanced in size, with the experimental and control groups having similar numbers of participants. However, the 4-week groups, especially the experimental group, have fewer participants than the 8-week groups, which could influence statistical power when comparing groups.

The study group was well-balanced by gender, but there were differences in the number of participants across groups and ages. These imbalances were considered when interpreting results, as smaller groups such as the 4-week experimental group or age 4 years 4 months old to 4 years 10 months old group might have lacked sufficient statistical power. The duration (4 vs. 8 weeks) might have influenced the outcomes, particularly since longer intervention was expected to provide more time for learning. Comparisons between these durations were considered in the following analyses. The age differences could have played a role in performance, as older children who were 4 years old may have had developmental advantages over younger ones who were 3 years old. When I analyzed the data, I considered these distributions to check for potential confounding factors or subgroup effects such as differences by gender, age, or group size. The teachers and the children were distributed across the groups as shown in Table 2.

Table 2

Descriptive Overview of Participating Teachers and Children by Group Type and Intervention Duration

Teachers pseudonym	Group	Duration	Number of children
Destiny	Experimental	4-week	15
Charlotte	Experimental	8-week	6
Gisele	Experimental	8-week	11
Klara	Experimental	8-week	12
Jane	Control	4-week	6
Victoria	Control	4-week	7
Darcie	Control	4-week	6
Agnes	Control	4-week	2
Grace	Control	8-week	28

A total of nine early childhood educators, all female, participated in the study. They were responsible for delivering the geometry program (or continuing their regular instruction in the control condition). Although one control group teacher had provided consent to participate in the study with her classroom, she delegated the children to another participating teacher colleague and two assistants during the scheduled observation periods, choosing instead to focus on administrative tasks outside the classroom. As a result, she did not participate in the interviews and observation notes, yet the children in her classroom joined Grace's group during the intervention. After outlining who participated, I turn to when and how data were gathered, specifying the sequence and cadence of observation, testing, and interviews.

Data Collection Procedure

After recruiting participating centres and securing consent from teachers, parents, and children (see Appendix A), I began classroom observations in all participating classrooms. These observations were conducted two times a week and were used to document pre-existing instructional practices in geometry. Based on these initial observations and discussions with administrators, classrooms were assigned to either the experimental or control group. To establish a baseline measure of children's shape recognition and abstraction skills, I administered the Geometric Shape Recognition Test (GSRT) individually to each child (see Appendix A).

Once group assignments were complete, I prepared experimental classrooms for the intervention by setting up a professional development space on Microsoft Teams and Google Drive. Teachers received sets of Froebel Gifts to be used with children during the intervention (20 sets in total). To support the teachers' understanding of the program and materials, I created and distributed 13 instructional videos. These included one video on the abstraction process, one introducing the Froebel approach, one outlining the FG-EGP, and 10 videos demonstrating the use of Froebel Gifts 1 through 10. Teachers in the experimental classrooms had 2 weeks to engage with the training materials, ask questions, and share reflections before the formal start of the intervention on April 9, 2024.

During the intervention, FG-EGP activities were implemented three times per week in the experimental classrooms. I continued visiting all classrooms throughout this period, documenting instructional strategies, children's use of vocabulary, and the integration of manipulatives. Control classrooms were also visited with the same frequency to observe their ongoing geometry instruction. Due to scheduling differences, four classrooms (one experimental and three control) followed a 4-week schedule, while the remaining four (two experimental and two control) completed the full 8-week intervention. This variation allowed for comparisons between short- and long-term program implementation. To evaluate how instructional components influenced children's abstraction, the GSRT was re-administered as a post-test following the intervention. Table 3 displays the data collection procedure in three phases. To make these procedures transparent, I then describe the instruments used to collect quantitative and qualitative data.

Table 3

Data Collection Procedure

Phase 1: Baseline observation (first round for 2 weeks) and pretest

Days	Groups						Timeline
Mon	Victoria	Jane	Darcie	Grace	Charlotte	Destiny	18–28 March 2024
Tue	Klara	Gisele	Destiny	Agnes			
Wed	Victoria	Jane	Darcie	Grace	Charlotte	Gisele	
Thu	Klara	Gisele	Destiny	Agnes	Jane		
Fri	Victoria	Darcie	Grace	Charlotte	Klara	Agnes	

Pretest with GSRT

Groups	No. of children	Dates
Destiny and Agnes	17	3, 4, 5 April 2024
Victoria, Jane, and Darcie	19	8, 9 April 2024
Grace and Charlotte	34	10, 11, 12 April 2024
Klara and Gisele	23	15, 16 April 2024

Phase 2: Teacher training and second round of observations + intervention

Teacher training schedule

Group	Froebel Gifts delivered	Training window
Destiny	April 5, 2024	April 9–19, 2024
Charlotte	April 12, 2024	April 16–26, 2024
Klara and Gisele	April 16, 2024	April 16–26, 2024

Second round of observations + intervention

Group	Days – Time	Duration
Destiny	Mon, Wed, Thu (13:00–13:40)	April 22 – May 16, 2024
Agnes	Mon, Wed, Thu (13:50–14:30)	
Victoria	Tue, Thu (9:00–9:40), one 11:15–11:55 slot (Mon/Tue/Fri rotation)	April 29 – May 24, 2024
Jane	Tue, Thu (9:45–10:25), one 11:15–11:55 slot (Mon/Tue/Fri rotation)	
Darcie	Tue, Thu (10:30–11:10), one 11:15–11:55 slot (Mon/Tue/Fri rotation)	
Grace	Mon, Wed, Fri (9:00–9:40)	May 6 – June 28, 2024
Charlotte	Mon, Wed, Fri (9:45–10:25)	
Klara	Tue, Fri (12:45–13:25), Wed (11:15–11:55, on overlap weeks) / Thu (12:45–13:25, all other weeks)	May 6 – June 28, 2024
Gisele	Tue, Fri (13:30–14:10), Wed (11:15–11:55, alternating with Klara on overlap weeks), Thu 13:30–14:10, all other weeks)	

Phase 3: Post-test and teacher interviews

Post-test with GSRT

Group	No. of children	Dates
Destiny and Agnes	17	May 22, 27, 29, June 4, 2024
Victoria, Jane, and Darcie	19	May 28, 30, June 4, 6, 2024
Grace and Charlotte	34	July 2, 3, 4, 2024
Gisele and Klara	23	July 2, 4, 5, 2024

Teacher interviews

Teachers	Dates
Destiny and Agnes	June 12, 2024
Victoria, Jane, and Darcie	June 11, 18, 2024
Grace and Charlotte	July 8, 9, 2024
Gisele and Klara	July 10, 2024

Data Collection Tools

This study utilized three main data collection tools: (a) the Geometric Shape Recognition Test (GSRT) developed by Aslan and Arnas (2007b), (b) a structured classroom observation form designed by me, and (c) a semi-structured interview protocol for post-intervention teacher interviews. These tools were selected to complement one another and to support the concurrent mixed methods research design by providing both quantitative and qualitative insights into children's abstraction processes in geometry learning. I begin with the quantitative tool, the GSRT, which operationalizes recognition accuracy and abstraction.

Quantitative Data: Geometric Shape Recognition Test (GSRT)

To measure children's abstraction processes in shape recognition, the GSRT was employed (see Appendix A). Developed by Aslan and Arnas (2007b), the GSRT is grounded in earlier research on geometric cognition in early childhood (Clements et al., 1999; Hannibal & Clements, 2010; Satlow & Newcombe, 1998). Designed for children aged 3 to 6, the test includes four classification tasks—triangle, rectangle, square, and circle—each printed on an A4 sheet and containing 12 shapes, including typical examples, atypical examples with non-defining attributes, palpable and impalpable distractors, and non-examples.

The GSRT was developed based on findings that young children often misclassify shapes when non-defining attributes such as aspect ratio, skewness, orientation, or size are altered. To account for this, each task includes typical items (e.g., U1 for triangle, D3 for rectangle, K1 for square, DA1 for circle), as well as atypical shapes and distractors (e.g., U2–U8 for triangle). Palpable distractors resemble the target shape but are incorrect, while impalpable distractors are clearly unrelated.

In the triangle task, items such as U7 and U2 assess skewness, while U4 (45°) and U5 (90°) assess orientation. The rectangle task tests the effect of aspect ratio, orientation, and size through D5 (1:8 ratio), D1 (90°), D2 (45°), and D4 (small size). The square task includes K2 (rotated), K3 (smaller size), and K4 (both rotated and smaller) to assess concept understanding. For the circle task, DA5 tests size variation, while DA2 and DA3 test edge thickness.

Children were individually administered the test in a quiet corner of their classrooms. The testing environment was controlled to avoid visual shape cues. I asked the child to mark all shapes that belonged to the target category in red and all shapes that did not belong in blue. Once marking was complete, I asked the child to justify each selection ("Why do you think this is a [shape]?" or "Why is this not a [shape]?") to understand their reasoning for including or excluding each shape. Their explanations were recorded in the GSRT registration form. Responses were categorized into three types: (a) visual responses, based on surface-level or familiar appearances; (b) property-based responses, referencing mathematical attributes such as number of sides, angles, or corners; and (c) "I don't know" responses.

For scoring, each correctly identified shape received a score of 1, while incorrect selections received a score of 0. Since each of the four shape categories (triangle, square, rectangle, and circle) included 12 items, the maximum possible recognition score on the GSRT was 48 points. Following each recognition task, children were also asked to explain their reasoning for including or excluding a shape. Reasoning responses were scored separately: a visual response received 1 point, a property-based response received 2 points, and an "I don't know" response received 0 points. Therefore, the maximum possible reasoning score was 96 points if a child provided property-based explanations for all 48 items. These dual scoring systems allowed for a clear distinction between accuracy of recognition and depth of abstraction in children's geometric reasoning.

The primary outcome variable was the number of property-based responses, considered a key indicator of abstraction. Although total scores reflected shape recognition accuracy, they did not alone indicate abstract thinking. Therefore, verbal responses were essential in identifying cognitive shifts from visual to conceptual classification. Further analysis focused on children's ability to recognize atypical shapes across the four shape categories. These data provided insight into the specific non-defining attributes (e.g., skewness, orientation, kurtosis, size, edge thickness) that influenced classification.

The test was administered both before and after the intervention as a pre-test and post-test. Each test session took approximately 25–30 minutes per child, and testing in each classroom was completed over two to four days, depending on children's attendance and classroom schedules. The pre-tests and post-tests were carried out across classrooms sequentially with the entire pre-test phase completed over a 2-week period. Because test scores alone cannot explain mechanisms, I complement the GSRT with systematic classroom observations before and during the intervention.

Qualitative Data: Classroom Observation Before and During the Intervention

Observation was employed in this study as a qualitative data collection method to capture the naturalistic instructional practices and contextual variables that influenced children's learning. As defined by Burns (1999), observation entails studying classroom interactions and events as they naturally occur, and in this study, it served as a complementary method to understand the influence of teacher practices on children's abstraction processes in geometry.

The purpose of conducting classroom observations was to gather supplementary information about the geometry learning environment in both control and experimental classrooms before and during the implementation of the FG-EGP. Specifically, observations aimed to document the types of activities teachers designed to support shape recognition, the vocabulary they used to describe geometric properties, and the type of manipulative they used in children's daily experiences. These data provided rich insight into how geometry was integrated into the play-based learning environment characteristic of these early childhood settings.

Observations were conducted three times a week for each classroom, for approximately 40 minutes per visit. The observation period before the intervention lasted 2 weeks. During the intervention, classroom visits continued for the full 8-week implementation period. I observed all nine classrooms, four experimental and four control groups, using a structured observation focus form (see Appendix A or Table 4) to guide data collection consistently. To minimize disruption and bias, I maintained a passive presence in the classroom, positioning myself in a location suggested by the teachers and refraining from initiating interactions with the children unless prompted by them.

Teachers introduced me to the children as a guest who was interested in watching their activities. Children were told not to interact with me so that their natural behaviour would not be influenced. Prior consent was obtained from all teachers, parents, and children involved in the observations.

The daycares involved in the study followed Ontario's Early Learning for Every Child Today (ELECT) framework, which emphasizes play-based learning. In these environments, mathematical concepts such as sorting, classifying, composing, de-composing, recognizing and naming shapes, exploring attributes (e.g., sides, corners, symmetry), and engaging with measurement vocabulary emerged organically through play and structured thematic units. Teachers often embedded geometry concepts into seasonal or thematic explorations, using real-life contexts to support early mathematical thinking.

Observational data focused on three central areas: instructional activities and guidance, vocabulary usage, and manipulative integration. The observation focus form was used to record both teacher and child actions in these domains. For example, in the domain of vocabulary, the form guided attention to whether teachers used precise geometric terms such as "rectangle," "triangle," or "vertex," and whether children adopted these terms in their descriptions. Similarly, for manipulatives, the form noted what materials were used (e.g., blocks, sticks, Froebel Gifts) and how children engaged with them to explore geometric properties. During each classroom observation, I took comprehensive handwritten field notes in real time, and later reviewed and categorized these notes onto structured observation forms under the headings of instruction, vocabulary, and manipulatives.

A summary of the structured observation criteria is provided in Table 4, which includes detailed guiding questions across the three focal areas: activities/instructions, vocabulary, and manipulatives.

Table 4

Observation Focus Form

Observation focus	Observation points for teachers	Observation points for children
Activities / instructions	• Types of geometry-related activities • Guidance on composing/de-composing shapes • Connections between geometry and general knowledge • Activities exploring similarities/ differences between shapes	• Types of activities children choose • Questions children ask • Connections made between geometry and other domains • Evidence of composing/de-composing shapes
Vocabulary	• Use of shape names and terms (e.g., triangle, square, rhombus) • Use of geometric property terms (e.g., side, angle, vertex) • Use of 3D shape terms (e.g., cube, cylinder) • Guidance on correct terminology use	• Conversations with peers/teachers • Descriptions of shape properties • Accuracy and clarity of vocabulary used • Differentiation between 2D and 3D terms
Manipulatives	• Provision of manipulatives to represent abstract concepts • Opportunities for children to explore shapes with manipulatives • Alignment of manipulatives with geometry goals	• Types of manipulatives used • Ways children represent shapes/components • Use of manipulatives for sides, edges, vertices

This structured and systematic approach to observation allowed for consistency across all classrooms and provided a comparative lens between experimental and control settings. The qualitative insights gained from the observation data offered a deeper understanding of the instructional and environmental elements that potentially shaped children's abstraction in geometric shape recognition.

The form (see Appendix A) guided data collection on several classroom features relevant to abstraction development, including the types of geometry activities presented, the vocabulary used by both teachers and children, the manipulatives employed in instruction, and children's responses and interactions with abstract concepts and materials. Observations were conducted prior to the intervention (baseline phase) and during the 8-week intervention period. In the pre-intervention phase, observations captured existing instructional practices, helping to contextualize pre-test performance and group assignment. During the intervention, the same form was used to document changes in classroom practices, particularly in the experimental classrooms where the FG-EGP was implemented. Observations took place 2 days a week in each classroom. The observation schedule prioritized the experimental classrooms on days when FG-EGP activities were conducted. The rich qualitative data from these observations were essential for interpreting changes in GSRT scores and for understanding the learning processes not captured by test data alone. To further illuminate teacher intentions behind observed practices, I add post-intervention interviews using a semi-structured protocol.

Qualitative Data: Teacher Interview Protocol

Following the post-test phase, semi-structured interviews (see Table B in Appendix A) were conducted with all nine participating teachers to collect additional qualitative data about their instructional strategies, experiences with the program, and perceived changes in children's learning. Each interview lasted approximately 45 to 60 minutes through Microsoft Teams or Zoom, depending on the teacher's availability and preference. Interviews were audio recorded with participants' consent and transcribed by me.

The interview protocol was structured around five central themes of geometry instruction: two-dimensional shapes, three-dimensional solids, sides/edges/lines, vertices, and angles. For each concept, teachers were asked about their instructional experiences, the vocabulary they and their students used, and how manipulatives were incorporated into learning activities. Interviews also probed how teachers scaffolded abstraction and corrected or clarified students' use of mathematical terms.

The purpose of the interviews was to contextualize the observed classroom practices and test outcomes by gaining insight into the reasoning behind teacher choices. Teachers' reflections on vocabulary usage, instructional goals, and challenges provided valuable triangulation for interpreting changes in children's GSRT scores. Having defined the measures, I next specify the quantitative analytic methods used to test intervention effects.

Quantitative Data Analysis

The quantitative phase of this study aimed to assess the impact of the FG-EGP on the abstraction processes of 3- to 4-year-old children in shape recognition, as measured by the Geometric Shape Recognition Test (GSRT). Data were collected using a pre-test and post-test design, and analyzed with a two-way repeated measures Analysis of Variance (ANOVA). All statistical analyses were conducted using SPSS.

The dependent variable was children's use of property-based responses in the GSRT, which served as an indicator of abstract thinking. The independent variable was the instructional condition: children either participated in the FG-EGP (experimental group) or continued with their usual curriculum (control group).

Table 5

Example of GSRT Coding

Code of the shape	Recognition of the shape scores (points): (Correct – 1) (Incorrect – 0)	Descriptions	Description scores (points): (Visual – 1) (Property – 2) (I don't know – 0)
UC1	1	It is not a triangle because the sides are curved or skewed.	Property – 2
U1	1	It is a triangle because it looks like a Christmas tree.	Visual – 1
DC1	1	It is not a rectangle because two sides are curved or skewed.	Property – 2
D2	1	It is a rectangle because it looks like a door.	Visual – 1
KC1	0	It is a square because it looks like a box.	Visual – 1
K1	1	It is a square because it has four sides, and they are the same length.	Property – 2
DAC4	0	It is a circle because it looks like that.	Visual – 1
DA1	1	It is a circle because it has one curved side.	Property – 2
Total	6		12

The repeated measures ANOVA was used to examine changes in children's GSRT performance over time (within-subjects factor: pre-test vs. post-test) and across instructional conditions (between-subjects factor: experimental vs. control group). The focus was on the interaction between time and group to determine whether the intervention led to significantly different gains in property-based reasoning.

Table 6

Children's Response Coding (Aslan & Arnas, 2007b)

Category	Verbal response	Description scores (points)
Visual	• Draws on paper or in air, saying "It looks like this", 'It looks like [doesn't look like] a [shape name]" • Reference to another shape on same page: "Same as this one" • Reference to another shape on same page: "This not the same as that" • Visual references to lines that are not horizontal or vertical: "It's slanty", "Pointy", or "It has corners" • "Looks like a [object name]" • Reference an object and saying "It doesn't look like a [object name]" • Skinny/fat/long • Big/small • Orientation • "That's the way it is" • "God does it like that" • More than one visual response	1
Property	• Round/curved/no straight sides/no corners • Number of angles • Number of sides • Type of lines • Length of sides • More than one property response	2
I don't know	• "I don't know"	0

Table 6 shows the categorization of children's responses to why they may or may not categorize the shown shape as a member of that shape class.

Each GSRT item was scored as correct (1) or incorrect (0), and verbal justifications were categorized as property-based (2), visual (1), or “I don’t know” (0). Table 5 illustrates sample coding of children’s GSRT responses. The detailed coding allowed for the identification of trends in abstract reasoning by measuring the frequency of property-based responses.

A power analysis using G*Power 3.1.9.7 confirmed that a minimum sample size of N = 54 would be sufficient to detect a medium effect (f = .25) with α = .05 and power = .95 for repeated measures ANOVA. The study sample (N = 93) was adequate for testing the proposed hypotheses. To contextualize those statistical results, I then outline the qualitative analytic approach used to derive themes from observations and interviews.

Qualitative Data Analysis

The qualitative data collection involved two rounds of structured classroom observations and post-intervention semi-structured interviews with teachers. This component of the research aimed to provide contextual understanding of instructional activities, vocabulary use, and manipulative mediation that might explain the quantitative outcomes.

Observations were conducted before and during the intervention to explore geometry learning experiences in both control and experimental groups. The observation protocol focused on: (a) teacher-led activities and instructions, (b) use of geometry-specific vocabulary by both teachers and children, and (c) the use and purpose of manipulatives in shaping children’s understanding of geometric concepts. Observational data were recorded using a structured observation form (see Table 4).

After the post-tests, I conducted one-on-one, semi-structured interviews with each actively participating teacher to gain insight into their instructional strategies, use of vocabulary, and their experience facilitating abstract geometry learning using manipulatives. The interviews were transcribed verbatim, and all qualitative data (interviews and field notes) were analyzed using thematic analysis, following guidelines from Evans (2018) and Maguire and Delahunt (2017).

Themes were generated through a systematic process of coding, categorizing, and synthesizing meaningful patterns related to geometry instruction. Three overarching themes guided the analysis:

- Instructional and Activity Design: This included teacher efforts to prompt abstraction through questioning, real-life analogies, shape comparison tasks, and guided inquiry.
- Manipulative Mediation: This theme captured how teachers selected and used physical materials to represent geometric concepts, as well as how they supported children in composing and decomposing shapes.
- Vocabulary and Geometric Discourse: This included teachers' and children's use of precise or informal terminology, and the ways teachers guided children toward more accurate mathematical language.

The qualitative data were analyzed using thematic analysis, following the six-step process outlined by Braun and Clarke (2006) and further elaborated by Maguire and Delahunt (2017). The analysis began with familiarization and initial coding of classroom observation notes. I first conducted a close reading of all observation data, manually colour-coding excerpts according to three overarching analytical categories that guided the study: (a) instructional activities, (b) vocabulary and discourse, and (c) manipulative use. This initial round of coding helped identify recurring instructional patterns and meaningful interactions between teachers and children that reflected how geometry was taught and experienced in the classroom.

Following this, I developed sub-categories under each of the three main categories to capture finer distinctions among instructional behaviours, language patterns, and manipulative use. During this process, several sub-categories overlapped or appeared redundant. To enhance analytic clarity, I used ChatGPT as an organizational tool to assist in grouping and labeling similar sub-categories. The AI-generated clusters served as a preliminary organizational framework, which I then critically reviewed by comparing them against the raw data. Through iterative verification, I refined, merged, or renamed sub-categories to ensure that they accurately represented the observed classroom practices.

Once a coherent and logically structured set of categories was established, I reviewed the coded data across sources, using first-round and second-round observations to confirm internal consistency and coherence within each main category. I continuously moved back and forth between the coded excerpts and the developing framework to ensure that the analysis remained grounded in the data rather than in preconceived assumptions.

The teacher interviews were analyzed using the same three overarching analytical categories established for the observations: Instructions and Activities, Vocabulary and Discourse, and Manipulatives. Since the interview questions were intentionally designed around these categories, a separate thematic analysis was not conducted for the interviews. The interview protocol itself (see Appendix A) clearly reflects these categories. Excerpts from the interviews were then selectively incorporated into the results sections to support, illustrate, or contrast the patterns identified in the observational data, thereby enriching the interpretation and triangulation of findings.

These qualitative findings were used to triangulate the quantitative data and provide a deeper understanding of the mechanisms through which the FG-EGP may have supported children’s abstraction, illuminating both statistical outcomes and pedagogical processes. The integrated analysis explored how the FG-EGP contributed to children's shift from visual to property-based reasoning in shape recognition and described the instructional context in which these shifts occurred. Finally, I address how the study's design and analyses were safeguarded against bias by detailing validity, reliability, and anticipated criticisms.

Validity, Reliability, and Criticism

To ensure the credibility and robustness of the research findings, careful attention was given to the validity and reliability of the study's instruments, procedures, and analyses. This section evaluates the methodological soundness of both the quantitative and qualitative components, discusses measures taken to enhance internal and external validity, and addresses potential limitations and criticisms that may affect the interpretation or generalizability of the results. Recognizing these strengths and weaknesses allows for a more balanced understanding of the study's contributions and areas for future improvement. I start with validity considerations, explaining how internal, construct, content, and ecological validity were supported.

Validity

To ensure internal validity, this study employed a quasi-experimental pretest-posttest control group design with multiple groups differentiated by intervention type (experimental vs. control) and duration (4 vs. 8 weeks). The inclusion of both pre- and post-tests using the validated Geometric Shape Recognition Test (GSRT) enabled the detection of changes in children's abstraction over time. Triangulation of data sources, including test scores, classroom observations, and teacher interviews, strengthened construct validity by allowing multiple forms of evidence to converge on the same phenomenon: abstraction in geometric shape recognition.

Content validity was addressed by using a test (GSRT) that explicitly includes a range of typical, atypical, and non-examples for each shape type, tapping into both recognition and abstraction. The test was grounded in prior developmental research on geometry (Aslan & Arnas, 2007b; Clements et al., 1999; Satlow & Newcombe, 1998), ensuring its theoretical appropriateness for children aged 3 to 4.

Ecological validity was supported by conducting observations and tests in children's natural classroom environments, without disrupting daily routines. This allowed findings to reflect authentic instructional contexts and children's everyday engagement with geometry.

In the qualitative component, validity was enhanced through prolonged engagement in the field (over 8 weeks), data triangulation (observations and interviews), and systematic thematic analysis. Member checking with the teachers were completed after I transcribed the interviews. However, member checking with the children was not feasible due to the young age of participants and the study's focus on observable instructional patterns rather than children-participant-reported experiences. Yet, clear audit trails in coding, detailed observation protocols, and teacher interviews helped preserve interpretive validity. I then turn to reliability, describing procedures that enhanced consistency in measurement and coding.

Reliability

Instrument reliability was addressed through the use of the GSRT, a previously published and structured assessment tool with clear scoring criteria. Each child's responses were scored for both recognition accuracy and explanation type (visual, property-based, or "I don't know"), reducing ambiguity in scoring.

For inter-rater reliability, the coding of children's verbal responses was conducted by me; however, all coding decisions were guided by established coding schemes (Aslan & Arnas, 2007a; Aslan & Arnas, 2007b), reducing subjectivity. Additionally, consistent procedures were followed across classrooms during data collection, aided by structured tools (e.g., observation focus forms and interviews with the teachers).

The qualitative data analysis adhered to systematic coding procedures, using predefined categories and thematic coding supported by literature. The use of a structured observation form enhanced the dependability of observational data by maintaining consistency in what was recorded across multiple classrooms and time periods.

Furthermore, descriptive validity was maintained by ensuring that all teacher quotes, classroom observations, and test responses were reported verbatim or summarized with fidelity. I avoided inference beyond the data and used clear categories to distinguish types of abstraction. With validity and reliability established, I conclude by acknowledging limitations and plausible critiques to frame the interpretation of findings.

Criticism and Limitations

Despite its methodological strengths, the study faces several limitations, which are discussed below.

- *Lack of Random Assignment*: Due to the constraints of working in existing daycare settings, classrooms were not randomly assigned to experimental or control groups. This introduces the possibility of selection bias, despite efforts to match classrooms based on demographic and institutional similarities.
- *Uneven Group Sizes*: The 4-week experimental group had a smaller sample size ($n = 15$) compared to the 8-week groups, which may have reduced the statistical power of comparisons involving this group and increased the margin of error in effect size estimation.
- *Researcher as Sole Observer*: While I used structured tools and remained non-intrusive, the use of a single observer may introduce unintentional bias in data interpretation, particularly in qualitative coding.

- *Potential Teacher Expectancy Effects*: Teachers in the experimental group were aware of the study's goals and received training through professional development, which may have influenced their instructional behaviour. Although this is an expected outcome of curriculum intervention research, it limits the ability to isolate the program's effects from teacher motivation and engagement.
- *Limited Generalizability*: The study was conducted in four daycare centres within a specific geographic region (Niagara, Ontario) and with a relatively small number of teachers (n = 9), all of whom were female. These contextual factors limit the external validity and generalizability of the findings to broader early childhood education settings or more diverse populations.
- *Age Variation*: Participants were 3 to 4 years and 10 months old, spanning nearly two years in developmental age. Although the majority were 3 years 9 months old to 4 years 3 months old, age-related cognitive differences may have influenced performance, especially given the rapid developmental changes that occur between ages 3 and 4.
- *Short Intervention Window*: The 4-week implementation period may have been insufficient to observe substantial shifts in abstraction for some children, especially in contrast to the 8-week group. Future studies might explore longer interventions or follow-up testing to assess retention.

Despite these limitations, the study provides valuable insights into how structured geometry instruction, particularly through manipulative use, teacher discourse, and comparison-based activities, can promote abstraction in early childhood mathematics. The integrated design allows for a nuanced understanding of both measurable outcomes and instructional processes.

With the research design, participants, instruments, procedures, and analytic strategies established, the dissertation now turns to the study’s findings. Chapter 4 presents the results of the embedded mixed-methods analysis, beginning with the quantitative GSRT outcomes that evaluate the impact of the FG-EGP on children’s abstraction in shape recognition across groups and durations. It then reports the qualitative findings from classroom observations and teacher interviews to illuminate how instructional activities, vocabulary, and manipulative use functioned as mechanisms that help explain the quantitative patterns.

CHAPTER FOUR: FINDINGS

Chapter 4 presents the findings from both the quantitative and qualitative strands of this mixed-methods study, which investigated the impact of the FG-EGP on 3- to 4-year-old children's abstraction processes in geometric shape recognition. All statistical analyses were conducted using SPSS. Quantitative results from pre- and post-test comparisons revealed that children in the experimental groups, particularly those in the 8-week intervention, demonstrated statistically significant improvements in both shape recognition and property-based verbal responses, compared to their control group peers. Statistical procedures including ANCOVA, independent and paired samples t-tests, Mann–Whitney U, and Wilcoxon Signed-Rank tests were used to assess change scores and group differences. The findings showed meaningful effect sizes and highlighted that extended and structured engagement with the FG-EGP supported children's abstract reasoning and geometric language development.

The qualitative findings complement the quantitative data by offering insight into how geometry teaching and learning unfolded across classrooms before, during, and after the intervention. Drawing on systematic classroom observations and semi-structured teacher interviews, the analysis revealed clear contrasts in instructional practices, vocabulary use, and manipulative-based engagement between control and experimental groups. In the experimental settings, especially those that implemented the full 8-week FG-EGP intervention, teachers employed intentional scaffolding, concept-rich language, structured activities with Froebel Gifts, and multi-sensory exploration to support abstraction. Peer collaboration and verbal reasoning were central to children's learning. In contrast, control group classrooms demonstrated more variable and often incidental geometry instruction with limited abstraction. All teacher and child names mentioned in this chapter are pseudonyms to protect participant confidentiality. Together, these findings provide robust evidence that the FG-EGP positively influenced both conceptual understanding and the use of mathematical vocabulary in early childhood geometry learning. To begin addressing the first research question, the following section presents the quantitative findings that establish the statistical foundation for understanding the impact of the FG-EGP intervention.

Quantitative Findings

In this section, I first present the demographic and descriptive statistics that delineate participant characteristics (age, gender, group assignment) and pre- to post-test score distributions, which provide essential context and highlight baseline differences. Next, I report the normality checks to confirm whether the data meet the assumptions for parametric testing. Following this, the chapter presents within-group comparisons using paired-samples t-tests and corresponding effect sizes, both overall and broken down by intervention condition (8- and 4-week experimental and control groups), to measure changes in scores over time. This is complemented by between-group analyses, where independent-samples t-tests (with variance checks via Levene's test) and effect sizes assess differences in performance across groups and verify baseline equivalence.

To account for baseline differences and variance heterogeneity, the chapter includes a ANCOVA, followed by Games–Howell post hoc comparisons, which provide adjusted mean comparisons across all study groups. Diagnostic checks, such as residual plots and normality tests were examined to ensure the tests used are robust and appropriately interpreted. Finally, because verbal response data were ordinal and included cells with low counts, Mann–Whitney U tests examine the impact of the FG-EGP intervention on children's abstraction and property-based shape recognition, rounding out a comprehensive statistical exploration that directly connects to the study's core questions: How did the differences in the geometry activities, manipulatives, and the vocabulary promoted by the teachers impact 3- to 4-year old children's abstract thinking skills in shape recognition?

The second research question serves as the overarching inquiry that frames the study. It explores how variations in instructional practices in the experimental and control groups, particularly those introduced through the FG-EGP, influence children's ability to abstract geometric properties during shape recognition tasks. This broad question is examined through three interrelated sub-questions, each focusing on a specific aspect of abstraction. Having outlined the scope of the quantitative analyses, the next subsection focuses specifically on whether and how the FG-EGP influenced children's abstraction processes in geometric thinking.

Exploring the Impact of the Froebel Gifts Early Geometry Program on Abstraction Processes in Early Geometric Thinking

The first sub-question investigates whether FG-EGP has a measurable impact on children's abstraction processes and, if so, in what ways this influence is observed. This sets the foundation for understanding the program's general effectiveness.

The following analysis investigates changes in participants GSRT scores in shape recognition across pretest and post-test measures, exploring differences by intervention duration (4- vs. 8-week programs) and group type (control vs. experimental). Participants, who were 3 years 3 months old to 4 years 10 months old, with a balanced gender split, were assigned to four groups, with the 8-week experimental group showing the most substantial gains in shape recognition. Overall, scores rose modestly but significantly, with the strongest improvement observed in the extended intervention group. Inferential statistics, including paired-sample t-tests, effect sizes, and ANCOVA controlling for baseline differences, are presented below, demonstrating that a longer, hands-on geometry intervention significantly boosts recognition performance, whereas a shorter version yields minimal impact.

Table 7 provides an essential overview of change and supports interpretation of all other results. Table 8 justifies parametric test use which is critical for understanding assumptions.

Table 7

Descriptive Statistics for Pretest and Post-test Scores by Group

Group	Measure	N	Mean	Std. Deviation
Control group (8 weeks)	Pretest scores	28	37.68	2.98
	Post-test scores	27	38.11	2.74
Experimental group (8 weeks)	Pretest scores	29	37.83	3.33
	Post-test scores	29	40.31	3.53
Control group (4 weeks)	Pretest scores	21	38.45	3.98
	Post-test scores	20	38.80	3.69
Experimental group (4 weeks)	Pretest scores	15	37.00	3.14
	Post-test scores	15	37.47	3.20
Total (all groups)	Pretest scores	93	37.80	3.31
	Post-test scores	91	38.86	3.42

The mean pretest score (M = 37.80, SD = 3.31) is slightly lower than the post-test score (M = 38.86, SD = 3.42). Skewness values suggest a moderately symmetrical distribution for both pretest (-0.475) and post-test (-0.236) scores. Kurtosis values indicate a near-normal distribution for both tests. The small increase in scores from pretest to post-test suggests potential intervention effects, though further statistical analysis is required to confirm this.

The experimental groups (8 weeks and 4 weeks) exhibit greater increases in post-test scores compared to the control groups. The largest improvement in scores is observed in the experimental group (8 weeks), with a mean increase of 2.48 points from pretest to post-test. The control group (4 weeks) shows minimal improvement (0.35 points), suggesting that intervention duration and type play a critical role in influencing outcomes. Longer intervention durations (8 weeks) may have caused more development in the recognition of geometric shapes than shorter ones (4 weeks), emphasizing the importance of duration.

Both pretest and post-test scores cluster around the mid-range (37–40), with a small percentage of participants scoring at the extreme ends (e.g., 27 or 47). The mode for post-test scores is 40, highlighting that many participants achieved higher scores after the intervention. Scores cluster in the mid-range, with the most frequent score being 39 (16.1% of participants). A few participants scored at the extreme ends (e.g., 27 or 45), suggesting some variability in pretest performance. The concentration of scores near the upper mid-range (35–39) reflects relatively uniform baseline abilities for most participants. For the pretest, scores ranged from 27 to 45, with the most frequent score being 39 (16.1% of participants). Other common scores included 36 (14.0%) and 38 (11.8%), indicating a clustering of scores around the upper-mid range. The post-test scores ranged from 31 to 47, with the most frequent score being 40 (15.1%), followed by 39 (11.8%) and 41 (12.9%). The shift in frequent scores from the pretest to the post-test suggests a general improvement in performance. The pretest and post-test mean scores were 37.80 and 38.86, respectively, with standard deviations of 3.312 and 3.418, showing similar variability in both distributions.

Overall, the mean increased slightly from 37.80 at pretest to 38.86 at post-test, and both distributions remained approximately normal, with similar variability (Table 1a, Appendix B). When examined by group, the 8-week experimental group demonstrated the clearest progress, with mean scores rising from 37.83 to 40.31 (Table 1b, Appendix B). Other groups showed only minimal gains, particularly the 4-week control group, whose pretest and post-test means remained virtually unchanged. This pattern suggests that extended exposure to the FG-EGP was associated with more substantial learning gains, whereas shorter participation produced little measurable improvement.

Table 8

Tests of Normality

	Groups	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Total score of pretest	Control group 8 weeks	.161	27	.069	.931	27	.075
	Experimental group 8 weeks	.188	29	.010	.908	29	.015
	Control group 4 weeks	.205	20	.027	.923	20	.115
	Experimental group 4 weeks	.158	15	.200	.948	15	.494
Total score of post-test	Control group 8 weeks	.146	27	.146	.925	27	.052
	Experimental group 8 weeks	.155	29	.074	.965	29	.423
	Control group 4 weeks	.177	20	.099	.929	20	.149
	Experimental group 4 weeks	.091	15	.200*	.987	15	.997

These group-level trends were further examined through assumption checks to ensure the appropriateness of subsequent analyses. Normality testing using the Kolmogorov–Smirnov and Shapiro–Wilk procedures indicated that most distributions approximated normality (Table 1c, Appendix B). Minor deviations were observed in the pretest data, particularly within the 8-week experimental group (*p* = .010, .015) and the 4-week control group (*p* = .027), while all post-test scores were nonsignificant (*p* > .05), confirming that the data were largely symmetrical and free from substantial skew. While some pretest scores deviate from normality in specific groups, post-test scores across all groups exhibit approximate normality. This supports the use of parametric tests like the paired samples t-test for post-test analysis but calls for caution when analyzing pretest data in groups where normality is not assumed (e.g., experimental group 8 weeks).

Preliminary inferential analyses further supported these descriptive trends. Across all participants, there was an average gain of about 1.03 points from pretest to post-test, reflecting a modest yet consistent improvement in geometric shape recognition (Table 2a, Appendix B). The strong positive correlation between pretest and post-test scores ($r = .926$, $p < .001$) indicated stable relative performance among children, suggesting that those who began with higher initial scores generally maintained their standing while still showing measurable growth (Table 2b, Appendix B). Together, these results confirm both a general upward trajectory in performance and strong internal consistency across assessments, providing a sound foundation for the subsequent *t*-test and ANCOVA analyses.

Table 9 provides the results of a paired samples t-test, comparing the pretest and post-test scores to determine whether the differences are statistically significant. This table includes all participants combined.

Table 9

Paired Samples Test (Paired Samples t-test for Within-Group Analysis)

	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			ds	Significance	
				Lower	Upper	t		One-sided p	Two-sided p
Total score of pretest - Total score of post-test	-1.033	1.303	.137	-1.304	-.762	-7.560	90	<.001	<.001

The paired samples t-test results reveal a statistically significant difference between the pretest and post-test scores ($t(90) = -7.560$, $p < 0.001$). The mean difference between the scores was -1.033, with a 95% confidence interval ranging from -1.304 to -0.762, indicating that post-test scores were consistently higher than pretest scores. The standard deviation of the differences was 1.303, and the standard error mean was 0.137, reflecting the precision of the mean difference estimate. These results confirm that the improvement in post-test scores is statistically significant and meaningful. The mean difference (−1.033-1.033−1.033) shows that post-test scores were, on average, higher than pretest scores. This statistically significant improvement suggests that the intervention had a positive impact on participants' performance.

Effect size analyses (Table 7j, Appendix B) further underscored the magnitude of this improvement. The observed difference corresponded to a large effect (Cohen's *d* = 1.30; Hedges' *g* = 1.31), indicating that the gains were not only statistically significant but also substantial in practical terms. This robust effect establishes a strong foundation for the subsequent group comparisons and ANCOVA models, confirming both the reliability and strength of the overall intervention effect.

Descriptive and inferential analyses for the 8-week control and experimental groups provide important context for interpreting the effects of the intervention. The 8-week control group exhibited only a marginal increase in mean scores from pretest (M = 37.81, SD = 2.95) to post-test (M = 38.11, SD = 2.74), with a statistically significant but small difference, *t*(26) = –2.530, *p* = .018 (Table 1b, Appendix B). The strong correlation between scores (r = .980, *p* < .001) suggests stable individual performance over time, likely reflecting natural development or test familiarity (Table 5f, Appendix B). In contrast, the 8-week experimental group showed a notable gain in performance, with mean scores increasing from 37.83 (SD = 3.33) to 40.31 (SD = 3.53) (Tables 1b, Appendix B). The correlation between pretest and post-test scores was also very strong (r = .939, *p* < .001), indicating consistent relative performance across participants

(Tables 9c, 51, Appendix B). These preliminary descriptive patterns set the stage for the inferential findings reported in Tables 11 and 12, which further confirm the statistically significant and educationally meaningful impact of the intervention.

Table 10

Paired Samples Test for Experimental Group 8 Weeks

	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			df	Significance	
				Lower	Upper	t		One-sided p	Two-sided p
Total score of pretest - total score of post-test	-2.483	1.214	.225	-2.944	-2.021	-11.017	28	<.001	<.001

Table 10 presents the results of a paired samples t-test assessing within-group differences in total scores between the pretest and post-test for participants in the 8-week experimental condition (N = 29). The analysis revealed a statistically significant improvement in performance following the intervention, $t(28) = -11.017$, $p < .001$ (two-tailed), with a mean difference of −2.483 (SD = 1.214, SE = 0.225). The 95% confidence interval for the difference ranged from −2.944 to −2.021, indicating a consistent and robust effect across the sample. These results provide strong evidence for the efficacy of the 8-week intervention, demonstrating a substantial enhancement in participants' scores, likely reflecting improved abilities in geometric shape recognition as a result of the FG-EGP program.

The strength of this improvement was further supported by the effect size analysis, which revealed a very large practical impact on children's performance; Cohen's $d = 1.21$; Hedges' $g = 1.25$ (Appendix B, Table 17). These values indicate that the observed gains were not only statistically significant but also educationally meaningful, reflecting substantial growth in children's geometric shape recognition. The confidence intervals for these estimates (Cohen's *d* CI = [–2.69, –1.39]) confirm the robustness and consistency of the effect across participants, further validating the efficacy of the full-duration, 8-week intervention.

Analyses of the 4-week conditions provided additional perspective on the role of intervention duration. In the 4-week control group, there was a slight increase in mean scores from pretest (M = 38.45, SD = 3.98) to post-test (M = 38.80, SD = 3.69) (Table 5a), yielding a statistically significant but small difference, $t(19) = –2.666$, $p = .015$ (Table 5c), with a moderate effect size (Cohen's $d = 0.587$) (Table 5d). The high correlation between pretest and post-test scores ($r = .991$, $p < .001$) (Table 5b) suggests strong consistency in individual performance, pointing to minimal gains likely due to developmental progression or test familiarity. In contrast, the 4-week experimental group showed a modest but statistically significant improvement, with scores rising from 37.00 (SD = 3.14) to 37.47 (SD = 3.20) (Table 22), $t(14) = –2.824$, $p = .014$ (Table 24). Although the mean difference was only 0.47 points, the effect size (Cohen's $d = 0.640$) and its confidence interval indicate that the observed gain is both statistically reliable and educationally meaningful (Table 5h). The strong correlation ($r = .980$, $p < .001$) between pre- and post-test scores in this group (Table 5f) also reflects consistent relative performance across participants. These findings suggest that even a brief implementation of the FG-EGP had a measurable impact on children's geometric shape recognition, outperforming natural gains observed in the control condition.

The mean difference of 0.47 between pretest and post-test scores in the 4-week experimental group represents a small but meaningful improvement in children's ability to correctly identify geometric shapes on the Geometric Shape Recognition Test (GSRT). Since each correctly recognized item within the four shape categories (triangle, rectangle, square, and circle) was scored as 1 point, a 0.47-point gain means that, on average, children correctly recognized about half an additional item after participating in the 4-week FG-EGP. While this may appear numerically small, it has practical significance in the context of early childhood geometry learning. Because most children already recognized the typical examples of each shape at the beginning of the study, any increase in score indicates that they began to extend their understanding to include atypical shapes (such as tilted or stretched versions) and to differentiate non-examples that share superficial similarities but lack defining properties.

Table 11

Independent Samples Test =Group (1 2) Comparison (Control vs. Experimental, 8 weeks)

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	T	Df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Total Score of Pretest	Equal variances assumed	.063	.803	-.178	55	.430	.860	-.149	.838	-1.828	1.530
	Equal variances not assumed			-.178	54.702	.430	.859	-.149	.836	-1.825	1.527
Total Score of Post-test	Equal variances assumed	.363	.549	-2.594	54	.006	.012	-2.199	.848	-3.899	-.499
	Equal variances not assumed			-2.617	52.348	.006	.012	-2.199	.840	-3.885	-.513

In other words, a 0.47-point increase suggests that through hands-on exploration and guided comparison activities, children started to generalize their geometric concepts beyond simple recognition and develop a more flexible, property-based understanding of shapes. Thus, even a fractional increase in the total score reflects an important qualitative shift in cognitive processing, not just a minor quantitative change.

A comparison between the 8-week control and experimental groups, as shown in Table 11 revealed that both began the study with nearly identical levels of geometric shape recognition. The mean difference in pretest scores was minimal (–0.15), and the independent samples *t*-test confirmed no statistically significant difference at baseline, *t*(55) = –0.18, *p* = .86 (two-tailed). The 95 % confidence interval for this difference (–1.83, 1.53) included zero, confirming that both groups were comparable before the intervention. Levene's test for equality of variances was nonsignificant (*F* = 0.06, *p* = .80), supporting the assumption of homogeneity and indicating proper randomization across the two groups.

Following the intervention, however, a clear divergence emerged. The 8-week experimental group significantly outperformed its control counterpart on the post-test, *t*(54) = –2.59, *p* = .012 (two-tailed), with a mean difference of –2.20 points (SE = 0.85; 95 % CI [–3.90, –0.50]). Levene's test again showed equal variances (*F* = 0.36, *p* = .55), confirming the robustness of the comparison. These results demonstrate that while both groups started from the same baseline, participation in the full-duration FG-EGP intervention led to a statistically and educationally significant improvement in geometric shape recognition.

Effect-size analyses supported this interpretation. The difference between the 8-week experimental and control groups corresponded to a medium-to-large effect; Cohen's *d* = 0.69 (Appendix B, Tables 6a–6d), highlighting that the gain was both statistically reliable and practically meaningful. When compared with the 4-week experimental group, the effect grew even stronger; *d* = 0.83 (Appendix B, Table 6g), underscoring the importance of intervention length. In contrast, the 4-week experimental group did not significantly outperform its control condition (*d* ≈ 0.38), showing that shorter exposure produced limited benefit. Descriptive results further reinforced this pattern: both experimental groups began with similar pretest means (M = 37.83 for 8-week; M = 37.00 for 4-week), yet only the 8-week group reached a post-test mean exceeding 40, confirming that extended engagement with the FG-EGP was critical for effect realization.

A direct comparison between the two experimental groups, those who received the FG-EGP for four versus eight weeks, further highlighted the influence of intervention duration on children's learning outcomes. Both groups began with comparable baseline performance (*t*(42) = 0.80, *p* = .43; 95 % CI [–1.27, 2.92]), confirming that initial abilities were equivalent. However, by the post-test stage, the 8-week group scored significantly higher than the 4-week group (*t*(42) = 2.61, *p* = .012), with a mean difference of approximately 2.84 points (95 % CI [0.65, 5.04]). The effect size for this difference was large; Cohen's *d* ≈ 0.83 (Appendix B, Table 18e), indicating that the longer duration produced not only statistically reliable but also educationally meaningful gains. In contrast, the shorter, 4-week implementation yielded only modest improvement and did not differ significantly from control conditions. These results reinforce that the FG-EGP's effectiveness depended strongly on sustained exposure. Children required extended engagement with the program to achieve measurable advances in geometric shape recognition.

The 2.84-point increase represents more than just statistical improvement; it indicates observable learning progress among the children who participated in the 8-week FG-EGP. At the start of the study, children primarily recognized typical shapes (e.g., a standard triangle or square), the most familiar and commonly presented examples. By the post-test, the children in the 8-week group began to recognize atypical versions of shapes (e.g., a stretched or rotated triangle such as the items U2, U3 in GSRT – see Appendix A), differentiate non-examples (figures that do not belong to a shape category, such as shapes with curved edges, items KC7, DC1 or UC2 in GSRT – see Appendix A), and apply geometric reasoning to identify defining properties (e.g., number of sides and vertices). Thus, the 2.84-point increase shows that children did not merely recall familiar forms but expanded their understanding of what counts as a valid geometric example. This shift reflects conceptual abstraction, the ability to generalize geometric properties across variations. The 2.84-point gain practically means that children moved from recognizing only typical shapes to identifying atypical shapes and correctly excluding non-examples, showing deeper and more abstract understanding of geometric categories.

The strength of the observed difference between the two experimental groups was further supported by effect size analyses, which confirmed that duration played a critical role in shaping outcomes. The post-test comparison revealed a large and educationally meaningful effect (Cohen's $d = 0.83$; Hedges' $g = 0.82$; Glass's $\Delta = 0.89$), with confidence intervals excluding zero, indicating that the extended, 8-week implementation of the FG-EGP produced substantially greater gains than the shorter 4-week version (Appendix B, Table 6g). In contrast, pretest effect sizes were negligible (Cohen's $d = 0.25$), reaffirming that the two experimental groups began with comparable baseline abilities. These findings provide strong evidence that it was the duration of exposure, rather than differences in initial skill level or program content, that accounted for the significant improvement in children's geometric abstraction. The large effect size (Cohen's d ≈ 0.83) supports that this improvement was educationally meaningful, not just a small numerical difference. It demonstrates that the longer, sustained exposure (8 weeks) allowed children to internalize and apply geometric concepts more flexibly and accurately.

Additional comparisons across control and experimental conditions further supported this interpretation. Across the four groups, the shorter, 4-week implementation consistently produced only minimal gains, confirming that reduced exposure time limited the impact of the program (Appendix B, Tables 7a–7i). For instance, when comparing the 8-week and 4-week control groups, both began and ended with similar scores (pretest: M = 37.68 vs. 38.48; post-test: M = 38.11 vs. 38.80), and the small effect sizes (Cohen's $d = –0.23$ for pretest, –0.20 for post-test) were negligible. This pattern indicates that simply extending the timeframe, without the structured FG-EGP instruction, did not result in measurable improvement.

When the 8-week experimental and 4-week control groups were compared, the longer intervention produced higher post-test means (M = 40.31 vs. 38.80), though this difference did not reach statistical significance ($t = 1.45$, $p = .155$; 95 % CI [–0.59, 3.61]) and corresponded to only a small effect (Cohen's $d = 0.42$). Similarly, comparisons between the 8-week control and 4-week experimental groups revealed no meaningful differences, with both pretest and post-test scores remaining close (Cohen's $d = 0.22$). Collectively, these analyses demonstrate that the 4-week experimental condition failed to outperform any control condition, and that significant learning gains were realized only through the extended, 8-week implementation of the program.

Before proceeding to the main ANCOVA analysis, assumption checks were performed to ensure the robustness of these findings. Levene's tests indicated a violation of the homogeneity of variances assumption ($p < .001$), prompting the use of the Games–Howell post hoc procedure for comparisons with unequal variances (Appendix B, Table 8a). Although the resulting values approached, but did not cross, conventional thresholds for statistical significance ($p = .052–.055$), they continued to show a consistent pattern favoring the 8-week experimental group, with differences approaching or exceeding two points across comparisons.

These results, together with the adjusted mean estimates and pairwise contrasts that follow, provide further confirmation that the extended intervention period produced the most substantial and reliable improvements in children's geometric shape recognition.

Table 12

Estimates – Dependent Variable: Total Score of Post-test

Groups	Mean	Std. Error	95% Confidence Interval	
			Lower bound	Upper bound
Control group 8 weeks	38.120[a]	.162	37.798	38.442
Experimental group 8 weeks	40.307[a]	.156	39.996	40.618
Control group 4 weeks	38.204[a]	.189	37.828	38.580
Experimental group 4 weeks	38.252[a]	.219	37.817	38.686

a. Covariates appearing in the model are evaluated at the following values: Total Score of Pretest = 37.82.

Table 12 displays the estimated marginal means for each group's post-test performance, averaged at the covariate value of the pre-test mean (M = 37.82). The key finding is that the 8-week experimental group achieved the highest adjusted post-test mean (M = 40.307, 95% CI [39.996, 40.618]), clearly outperforming all other groups. The 4-week experimental group's adjusted mean was M = 38.252 (95% CI [37.817, 38.686]), which overlaps with the control groups' means (M = 38.120 and 38.204), suggesting no clear advantage over the control conditions.

Moving to Table 13 (Pairwise Comparisons), each mean difference represents a planned post hoc contrast, with statistical control for unequal variances via the Games–Howell method (necessitated by the significant Levene's test).

Table 13

Pairwise Comparisons – Dependent Variable: Total Score of Post-test

(I) Group	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference[b]	
					Lower bound	Upper bound
Control group 8 weeks	Experimental group 8 weeks	-2.187	.225	<.001	-2.795	-1.579
	Control group 4 weeks	-.084	.249	1.000	-.756	.589
	Experimental group 4 weeks	-.132	.272	1.000	-.866	.603
Experimental group 8 weeks	Control group 8 weeks	2.187	.225	<.001	1.579	2.795
	Control group 4 weeks	2.103	.245	<.001	1.441	2.766
	Experimental group 4 weeks	2.055	.269	<.001	1.330	2.781
Control group 4 weeks	Control group 8 weeks	.084	.249	1.000	-.589	.756
	Experimental group 8 weeks	-2.103	.245	<.001	-2.766	-1.441
	Experimental group 4 weeks	-.048	.290	1.000	-.832	.736
Experimental group 4 weeks	Control group 8 weeks	.132	.272	1.000	-.603	.866
	Experimental group 8 weeks	-2.055	.269	<.001	-2.781	-1.330
	Control group 4 weeks	.048	.290	1.000	-.736	.832

The results, as shown in Table 13, confirm that the 8-week experimental group significantly outperformed both control groups (difference ≈ 2.187 points, $p < .001$) as well as the 4-week experimental group (difference ≈ 2.055, $p < .001$). No other pairwise comparisons reached significance, indicating that the intervention effect was specific to the longer-duration experimental group. Together, these tables provide robust evidence that, after accounting for pre-test scores, the 8-week intervention had a significantly positive impact on post-test performance, unlike the shorter intervention or control conditions.

Further analyses reinforced this conclusion. The overall ANCOVA model confirmed a highly significant effect of group membership on post-test performance ($F(3, 83) = 40.97$, $p < .001$), while pretest scores emerged as the strongest predictor of post-test outcomes ($F = 1122.59$, $p < .001$), highlighting the importance of controlling for baseline ability (Appendix B, Tables 8d and 51). Visual inspection of the data supported these results: scatterplots showed a strong linear relationship between pre- and post-test scores, confirming test reliability, and group-size checks verified adequate statistical power (Appendix B, Tables 9a–9b).

Adjusted post-test means obtained through ANCOVA further clarified the pattern. The 8-week experimental group achieved the highest adjusted mean (M = 40.31, 95 % CI [39.99, 40.62]), distinctly higher than any other group. The 4-week experimental group (M = 38.29, 95% CI [37.84, 38.74]) overlapped substantially with both the 8-week control (M = 38.12, CI [37.80, 38.44]) and 4-week control (M = 38.22, CI [37.84, 38.61]), confirming that the shorter intervention did not lead to measurable improvement (Appendix B, Tables 9d–9e).

Diagnostic checks confirmed the robustness of these findings. A univariate ANOVA verified the overall group effect ($F(3, 83) = 40.97$, $p < .001$), validating the decision to conduct pairwise contrasts. Examination of residuals showed a mean near zero and moderate variability (~0.82), indicating no systematic prediction bias. Skewness values between −0.50 and −0.61 and kurtosis between 1.62 and 1.91 suggested a slightly left-skewed but acceptably peaked distribution, consistent with good model fit (Appendix B, Tables 10a–10b). Formal normality tests (Kolmogorov–Smirnov and Shapiro–Wilk) were significant ($p < .001$), implying minor deviations from perfect normality; however, given the sample size (≈ 90), ANCOVA is robust to such departures, and the validity of the conclusions remains intact (Appendix B, Table 10c).

Taken together, these results confirm that duration was the decisive factor in determining the program's effectiveness. Only the 8-week implementation produced statistically significant and educationally meaningful improvements in children's geometric abstraction, while the shorter 4-week version—whether experimental or control—did not yield measurable benefits beyond what could be achieved through standard practice. Sustained exposure to structured, manipulative-based activities therefore appears essential for fostering lasting conceptual growth in young learners.

Triangle Recognition

Triangle items were grouped into non-examples (UC1–UC5), typical (U1), and atypical forms based on skewness (U2, U7), orientation (U4, U5), and kurtosis (U3, U6). While no statistically significant group differences emerged, descriptive patterns revealed that the 8-week experimental group outperformed others in recognizing triangles with skewed and kurtosis-based deviations. Although the results did not reach significance, the mean gains suggest a trend that extended instructional exposure may support abstraction in identifying geometric features that deviate from familiar forms.

Table 14

Mean Gains in Children's Recognition of Skewed, Kurtotic, and Differently Oriented Triangles by Group and Duration

Group	Skewness Gain	Kurtosis Gain	Orientation Gain
Control group (8 weeks)	–0.037	+0.074	+0.150
Experimental group (8 weeks)	+0.103	+0.172	0.000
Control group (4 weeks)	–0.100	–0.100	0.000
Experimental group (4 weeks)	+0.200	+0.133	0.000

These gains as presented in Table 14, while not statistically significant, suggest that structured exposure through the FG-EGP may have supported the development of more generalized and property-based triangle recognition. Because each triangle item was scored dichotomously (1 = correct recognition, 0 = not recognized), a kurtosis gain of 0.172 represents approximately a 17.2 percentage-point increase in correct recognition of kurtosis-variant triangle items at the group mean level from pretest to post-test.

Rectangle Recognition

Rectangle items included non-examples (DC1–DC7), a typical example (D3), and atypical shapes varying by skewness (D4), orientation (D1, D5), and kurtosis (D2).

Table 15

Mean Gains in Children's Recognition of Skewed, Kurtotic, and Differently Oriented Rectangles by Group and Duration

Group	Skewness Gain	Kurtosis Gain	Orientation Gain
Control group (8 weeks)	–0.040	–0.037	0.000
Experimental group (8 weeks)	+0.137	+0.172	+0.214
Control group (4 weeks)	–0.050	0.000	+0.147
Experimental group (4 weeks)	+0.133	+0.133	+0.267

A one-way ANOVA revealed a statistically significant difference in recognizing skewed rectangles across the four groups, $F(3, 87) = 2.98$, $p = .036$, with a medium effect size ($\eta^2 = .093$). Although post hoc comparisons were not significant, descriptive data indicated higher mean gains among experimental groups. Recognition of kurtosis approached significance, $F(3, 87) = 2.53$, $p = .063$ ($\eta^2 = .080$), again with the 8-week experimental group showing the most improvement. Recognition of unusually oriented rectangles did not differ significantly across groups (F(3, 87) = 1.18, p = .324), though descriptive differences again favored experimental groups as presented in Table 15.

Square Recognition

Square items included non-examples (KC1–KC8), one typical example (K1), and atypical forms targeting orientation (K2, K4) and size (K3, K4). The ANOVA for size recognition was marginally significant, F(3, 87) = 2.53, p = .063, with a small-to-medium effect size ($\eta^2 = .080$). Orientation differences were not significant, F(3, 87) = 0.76, p = .521. However, descriptively, the 8-week experimental group had the highest mean gain for size as shown in Table 16, while the 4-week experimental group showed parallel improvement in both size and orientation.

Table 16

Mean Gains in Children's Recognition of Differently Sized and Oriented Squares by Group and Duration

Group	Size Gain (Kurtosis)	Orientation Gain
Control group (8 weeks)	–0.04	+0.19
Experimental group (8 weeks)	+0.17	+0.03
Control group (4 weeks)	0.00	+0.10
Experimental group (4 weeks)	+0.13	+0.13

Although these changes were not statistically significant, they are directionally consistent with the program's emphasis on guiding children toward identifying geometric properties across variable representations. My findings on square recognition, especially regarding children's responses to size (kurtosis) and orientation, align meaningfully with the study by Halat and Yeşil Dağlı (2016) in both content and implication. Halat and Yeşil Dağlı (2016) found that many preschool children struggled to recognize rotated (atypical orientation) squares, with 79% and

56% failing two rotated-square identification tasks. This aligns with my result that orientation differences were not statistically significant (F = 0.76, *p* = .521). However, my 4-week experimental group showed equal improvement in both size and orientation, suggesting some early benefit of structured exposure, even if the gains weren't statistically robust.

In contrast, the marginally significant ANOVA result for size-based square recognition in my study (F = 2.53, *p* = .063) supports Halat and Yeşil Dağlı's observation that children more easily recognize squares when they maintain a standard size and orientation. My finding that the 8-week experimental group had the highest mean gain in size-based recognition (+0.17) suggests that sustained, guided exposure through Froebel Gifts, composition and comparison tasks, helped children move beyond reliance on prototypical examples. After examining overall performance and shape recognition outcomes, the following subsection narrows the focus to changes in children's property-based verbal responses as indicators of abstract reasoning.

Changes in the Frequency of Property-Based Shape Descriptions Following the FG-EGP Intervention

The second sub-question narrows the focus by exploring whether FG-EGP affects the frequency of children's property-based responses, which serve as indicators of their abstract reasoning. To investigate whether FG-EGP influenced children's ability to abstract and apply geometric properties when recognizing shapes, several statistical analyses were conducted focusing specifically on verbal responses coded during the Geometric Shape Recognition Test (GSRT).

To investigate whether the FG-EGP influenced the frequency of children's property-based reasoning in geometric shape recognition tasks, two variables were computed as PropertyCountPre and PropertyCountPost. These variables represent the number of times each child gave a property-based verbal response (coded as “2”) during the pretest and post-test, respectively, across 48 shape items. Unlike the total verbal response scores which combine all response types including visual-based (coded as “1”) and “I don’t know” (coded as “0”), these variables specifically isolate and quantify the child’s use of geometric properties in justifying shape recognition.

Table 17

Wilcoxon Signed-Rank Test for Change in Property-Based Responses from Pretest to Post-Test

Ranks	N	Mean Rank	Sum of Ranks
Negative Ranks: Post-test < Pretest[a]	30	40.75	1222.50
Positive Ranks: Post-test > Pretest[b]	40	31.56	1262.50
Ties: Post-test = Pretest[c]	21	—	—
Total	91		
Z		–0.118	
Asymptotic Significance (2-tailed)		.906	

Note. Wilcoxon Signed-Rank Test was used to assess whether the frequency of property-based responses (coded as “2”) increased from pretest to post-test.

[a] Number of times each child gave a “2” in the post-test < pretest.

[b] Number of times each child gave a “2” in the post-test > pretest.

[c] Number of times each child gave a “2” in the post-test = pretest.

To assess whether there was a significant increase in property-based responses following the FG-EGP intervention, a Wilcoxon Signed-Rank Test was conducted. This non-parametric test was selected because the data consist of count values bounded between 0 and 48, and the assumption of normality may not hold. A statistically significant result would indicate that the frequency of children’s property-based reasoning increased from pretest to post-test, supporting the hypothesis that FG-EGP foster the development of abstract geometric thinking in early childhood.

Although a larger number of children showed an increase (n = 40) than a decrease (n = 30) in property-based shape recognition responses following the FG-EGP intervention, the Wilcoxon Signed-Rank Test as presented in Table 17 revealed that this difference was not statistically significant, Z = –.118, p = .906. These findings suggest that, overall, children did not significantly increase their use of geometric properties in verbal justifications across the full sample.

To corroborate these findings using a parametric method, a one-way between-groups ANOVA was also conducted. This decision was based on the observation that the distribution of property-based response frequencies showed slight deviations from normality, but not to a degree that would invalidate parametric assumptions entirely. The ANOVA, which assumes normally distributed data and homogeneity of variance, also revealed no statistically significant group differences.

Table 18

One-Way ANOVA Comparing Post-Test Property-Based Response Frequencies Across Four Groups

Source	SS	df	MS	F	p
Between groups	21.22	3	7.07	1.01	.392
Within groups	609.24	87	7.00		
Total	630.46	90			

These converging results across both parametric and non-parametric approaches strengthen the conclusion that the FG-EGP intervention did not produce statistically significant differences in children's frequency of property-based reasoning across the four groups. However, qualitative differences in gesture-based reasoning and subgroup trends may offer additional insights into developing abstraction, which are further explored in the qualitative section.

Children's verbal responses were carefully coded to differentiate between property-based recognition and visual-based recognition of geometric shapes. Specifically, responses were coded as property responses when children clearly referenced geometric properties, using vocabulary such as "three sides," "three corners," or "four sides," explicitly indicating awareness of the defining features of the shapes. For example, a child stating, "It is a triangle because it has three sides," was coded as a property response. There were some gesture responses during the pre-test, but they were very few and often unclear, mainly limited to a few children in one of the 4-week control classrooms, who were also slightly older than the rest of the sample. Because these gestures were sporadic and not yet a salient pattern, I did not initially code them as "gesture responses." They were instead captured within the broader "visual response" category since my initial coding focused primarily on verbal indicators of property-based reasoning. However, by the post-test, I observed that a substantial number of children in the experimental groups, especially the 8-week intervention, were consistently using gestures such as tracing edges, touching corners, or forming shapes with their hands to express geometric understanding. This emerging and repeated behaviour prompted me to recognize gesture as a distinct and meaningful response type and to begin coding these instances systematically as gesture responses rather than direct property-based responses. In short, gestures did occur at the pre-test but were too infrequent and indistinct to be coded separately at that stage. Their frequency, clarity, and conceptual relevance became apparent by the post-test, leading to the coding refinement and the observed statistical significance in gesture change scores.

Group Comparisons of Verbal Responses

Although a Chi-square test was conducted to examine the distribution of response types (see Table 11a in Appendix B), no statistically significant differences were found at either time point (pre-test p = .433; post-test p = .174). However, this test compared group-level distributions and not within-subject changes, which are more aligned with the theoretical focus of this research. Further issues with expected cell counts led to the use of Mann–Whitney U tests for more robust analysis.

Baseline equivalence between the 8-week experimental and control groups was confirmed using Mann–Whitney U (Tables 11b–11c; Appendix B). At pre-test, both groups performed similarly in verbal response scores, validating that any post-test differences could be attributed to the intervention.

At post-test, however, the 8-week experimental group significantly outperformed the 8-week control group, as can be seen below:

Table 19

Ranks – Total Verbal Response Score of Post-test

Experimental or control group	N	Mean Rank	Sum of Ranks
Control Group 8 weeks	27	24.02	648.50
Experimental Group 8 weeks	29	32.67	947.50
Total	56		

The Mann-Whitney Ranks table (Table 19) compares the total verbal response scores between the 8-week control group and the 8-week experimental group at the post-test stage. The experimental group (M = 32.67, Sum of Ranks = 947.50) had a higher mean rank than the control group (M = 24.02, Sum of Ranks = 648.50), suggesting that participants in the experimental group performed better in verbal responses after the intervention. However, to determine whether this difference is statistically significant, I need to examine the Mann-Whitney U test results and p-values.

The Mann-Whitney U test results (Table 20) indicate a statistically significant difference in total verbal response scores between the 8-week experimental and control groups at the post-test stage (U = 270.500, Z = -1.991, p = .046).

Table 20

Test Statistics

	Total Verbal Response Score of Post-test
Mann-Whitney U	270.500
Wilcoxon W	648.500
Z	-1.991
Asymp. Sig. (2-tailed)	.046

Note. Grouping variable: experimental or control group.

Since p < .05, this suggests that the experimental group outperformed the control group in verbal responses after the intervention, implying that the intervention had a meaningful impact on improving verbal response scores. The difference in ranks observed in Table 17 is now confirmed to be statistically significant, supporting the conclusion that the experimental program contributed to better verbal response performance.

The results presented in Tables 16a through 16d (see Appendix B) shed important light on why the 4-week experimental intervention failed to narrow performance differences in verbal responses. At pre-test (Tables 16a–16b in Appendix B), the 4-week control group outperformed its experimental counterpart significantly (U = 48.500, p < .001). This same pattern held at post-test (see tables 16b–c in Appendix B; U = 48.500, p < .001), indicating that the experimental condition did not close the initial gap.

Crucially, this disparity appears to stem from participant composition since the 4-week control group included older children who were 4 years to 4 years and 10 months old whose teachers emphasized mathematics and geometry activities, an instructional focus that may have boosted their verbal response skills beyond those of the experimental group. As a result, the 4-week experimental condition failed to match their performance, despite the intervention. Before the intervention, the two experimental groups (4-week vs. 8-week) were not significantly different in property-based responses (Tables 16d–16e; Appendix B). However, at post-test, the longer-duration group performed significantly better.

Table 21

Mean and Sum of Ranks for Experimental Groups on Post-test Verbal Responses

Group	N	Mean Rank	Sum of Ranks
Experimental group 8 weeks	29	25.67	744.50
Experimental group 4 weeks	15	16.37	245.50
Total	44		

Table 21 presents the distribution of post-test verbal response scores for the two experimental groups (8-week vs. 4-week). Each group's total verbal scores are ranked, with the 8-week group (n = 29) receiving a higher mean rank of 25.67, compared to the 4-week group's mean rank of 16.37 (n = 15). This suggests that, overall, participants in the longer intervention performed better in verbal responses than those in the shorter version. These ranks form the basis for the Mann–Whitney U-test (reported in Tables 68 and 69 in Appendix B) to evaluate whether this observed difference is statistically meaningful.

Table 22

Inferential Statistics for Post-test Verbal Responses: Comparison of Experimental Groups

	Total Verbal Response Score of Post-test
Mann-Whitney U	125.500
Wilcoxon W	245.500
Z	-2.281
Asymp. Sig. (2-tailed)	.023

Note. Grouping variable: experimental or control group.

Table 22 reports the results of the Mann–Whitney U test comparing post-test verbal response scores between the two experimental groups (8-week vs. 4-week). With a U value of 125.5, Wilcoxon W of 245.5, and Z = –2.281, the test yields an asymptotic two-tailed p-value of .023, which is below the conventional .05 threshold. This indicates a statistically significant difference in verbal response outcomes: participants in the 8-week intervention clearly outperformed those in the 4-week one. In other words, the longer program led to greater improvements in verbal responses, confirming that the duration of intervention had a meaningful impact on this outcome.

At baseline, the 4-week control group had a significantly higher mean rank in verbal response scores (Mean Rank = 32.88) than the 8-week control group (Mean Rank = 19.09), as confirmed by a Mann–Whitney U = 128.50, Z = –3.348, p < .001 (Table 18b in Appendix B). This indicates that the control groups did not start at equivalent levels, which is important when interpreting post-test comparisons. After the intervention period, the 4-week control group continued to outperform the 8-week control group (Mean Ranks: 31.25 vs. 18.63), with a Mann–Whitney U = 125.00, Z = –3.129, p = .002 (Table 18c – 18d in Appendix B). This persistence of differences signals that duration in control conditions alone did not equalize performance levels across baseline disparities.

Finally, baseline and post-test comparisons between the two control groups (Tables 18a–18d; Appendix B) indicated that the 8-week control group consistently outperformed the 4-week control group. This trend likely reflects differences in group composition, instructional context, or teacher emphasis on geometry, rather than effects solely related to the intervention itself. These variations highlight the influence of contextual and pedagogical factors on children's outcomes and, while informative, remain supplementary to the main analysis focused on the intervention effects. Having examined how the FG-EGP influenced children's overall abstraction and property-based verbal reasoning, the next subsection delves deeper into how children used attributes and gestures to define and describe geometric shapes, reflecting embodied dimensions of abstract thought.

Evaluating the Impact of the Froebel Gifts Early Geometry Program on Children's Property-Based Reasoning in Shape Recognition

Finally, the third and last sub-question examines the quality and nature of children's responses by asking how they use attributes and properties to define shapes, and whether this influence is positive or negative. Together, the three sub-questions provide a layered analysis of the overarching research question, progressing from general impact to specific cognitive behaviours. This structure offers a comprehensive view of how structured geometry instruction can support the development of abstract thinking in early childhood.

To evaluate whether the FG-EGP influenced children's use of geometric attributes and properties, gesture responses were analyzed as a key indicator of embodied abstraction.

Children's changes in gesture responses from pretest to post-test were examined descriptively and tested for normality (see Tables 19a and19b in Appendix B).

The mean change in gesture responses was 2.41 (SD = 3.41), with a 95% confidence interval ranging from 1.70 to 3.12. The mean change of 2.41 indicates that, on average, children produced approximately two to three more gesture-based responses in the post-test than in the pre-test. In other words, most children demonstrated a modest increase in using gestures, such as tracing edges, touching corners, or forming shapes with their hands, to identify or describe geometric figures. While the median change was +1, suggesting that at least half of the children showed small improvements, a few children exhibited much larger gains (up to 15 additional gestures). The distribution of gesture change scores was positively skewed (skewness = 1.49, SE = .25) and leptokurtic (kurtosis = 1.64, SE = .50), suggesting a concentration of lower scores with a few extreme increases. This pattern explains the positive skew in the data and suggests that although gesture use increased broadly, the most substantial improvements occurred among children in the 8-week experimental group, where the FG-EGP had the strongest impact on embodied abstraction.

Normality tests confirmed that the gesture change data deviated significantly from a normal distribution, as indicated by the Kolmogorov–Smirnov test (D(91) = .253, $p < .001$) and Shapiro–Wilk test (W(91) = .760, $p < .001$). Based on this non-normality, a non-parametric Kruskal–Wallis H test was conducted to compare gesture response changes across groups.

The Kruskal–Wallis H test in Table 23 indicated a statistically significant difference in gesture count change scores across groups, $\chi^2(3, N = 91) = 52.05$, $p < .001$. Mean ranks showed that the 8-week Experimental Group (M rank = 70.86) had the highest increase in gesture responses, followed by the 4-week Experimental Group (M rank = 49.63). In contrast, the 8-

week Control Group (M rank = 34.31) and 4-week Control Group (M rank = 23.00) showed smaller gains in gesture-based responses.

Table 23

Kruskal-Wallis H Test for Gesture Score Change Analysis

Group	N	Mean Rank
Control group 8 weeks	27	34.31
Experimental group 8 weeks	29	70.86
Control group 4 weeks	20	23.00
Experimental group 4 weeks	15	49.63

Note. Kruskal-Wallis H(3) = 52.05, p < .001.

Because this test revealed significant differences across groups, pairwise comparisons using Mann–Whitney U tests were conducted with a Bonferroni-corrected significance threshold of *p* < .0083. These analyses clarified which specific group differences accounted for the overall effect (see Table 24).

Table 24

Post Hoc Analysis

Group Comparison	U	Z	p-value	Mean Rank 1	Mean Rank 2	Interpretation
Control 8w vs experimental 8w	65.00	-5.46	< .001	16.41 (Control)	39.76 (Experimental)	Experimental 8w > control 8w
Experimental 8w vs experimental 4w	83.00	-3.35	< .001	27.14 (8w)	13.53 (4w)	Experimental 8w > experimental 4w
Control 4w vs experimental 4w	50.00	-4.18	< .001	13.00 (Control)	24.67 (Experimental)	Experimental 4w > control 4w

In Table 24, follow-up Mann-Whitney U tests with a Bonferroni-corrected significance threshold ($p < .0083$) results showed that the Experimental Group 8 weeks demonstrated significantly greater increases in gesture responses compared to the Control Group 8 weeks, $U = 65.00$, $Z = -5.46$, $p < .001$. Similarly, the Experimental Group 8 weeks showed significantly greater gesture gains than the Experimental Group 4 weeks, $U = 83.00$, $Z = -3.35$, $p < .001$. Additionally, the Experimental Group 4 weeks exhibited significantly greater increases in gesture responses compared to the Control Group 4 weeks, $U = 50.00$, $Z = -4.18$, $p < .001$. These findings clearly indicate that the FG-EGP intervention led to significant improvements in children's gesture-based shape responses, with the strongest effects observed in the 8-week implementation. This suggests that the program not only enhanced verbal abstraction but also fostered embodied forms of expression, such as gesture, as children increasingly used physical actions to represent geometric properties.

The findings presented in this section provide strong evidence that the FG-EGP positively influenced how children use attributes and properties to define shapes. Quantitative analyses revealed that gesture-based responses significantly increased, especially in experimental groups and more so with longer exposure. Qualitative insights further confirmed that hands-on, bodily engagement with shapes supported children's understanding of abstract geometric properties. Taken together, these results highlight that FG-EGP promotes not just verbal abstraction, but also embodied forms of abstract reasoning, marking a meaningful shift in how young children come to understand the abstract properties of shapes. While the quantitative findings establish clear evidence of growth in children's abstract reasoning and gesture-based understanding, the following section turns to the qualitative data to explore *how* these developments emerged within real classroom contexts through teachers' practices and children's interactions.

Qualitative Findings

This section presents the qualitative findings drawn from classroom observations and semi-structured teacher interviews, with a particular focus on how young children engaged with geometry-related experiences in their early learning environments before, during, and after the implementation of the FG-EGP.

The upcoming sections centre on two key sources of qualitative evidence that illuminate classroom dynamics and instructional practice. The first part presents themes from classroom observations, drawing on systematic notes taken before and during the FG-EGP intervention to identify patterns in children's engagement, reasoning, and use of manipulatives during geometry activities. The second part explores themes from teacher interviews, synthesizing insights from semi-structured conversations with educators about their instructional strategies, scaffolding techniques, vocabulary use, and perceptions of students' abstract understanding of shape properties. Together, these thematic analyses enrich and contextualize the quantitative outcomes, revealing *how* and *why* FG-EGP may influence children's abstraction processes, not just *whether* it does. To contextualize these outcomes, the first round of classroom observations conducted before the intervention offers a baseline view of how young children engaged with geometry concepts in their everyday learning environments.

First Round of Observations (Before the Intervention)

The observation data were collected from four daycare centres of the 180 licensed centres in the Niagara Region prior to and during the implementation of the intervention. The aim of the initial observations was to gain insight into the existing learning experiences and environments, serving as a baseline for understanding how children engaged with early mathematical ideas in their daily routines.

Table 25

Emerging Themes and Subthemes from the Initial Observations

Themes	Subthemes (codes)
Mathematical concept development	• Recognizing patterns in shapes and objects in their daily environment • Understanding mathematical concepts such as counting, measuring, size and height
Hands-on exploration and manipulatives	• Stacking, arranging, classifying and organizing materials • Using manipulatives to build mathematical and spatial awareness • Engaging in various kinds of experiential learning through physical objects
Inquiry-based learning and problem-solving	• Encouraging the children's curiosity with open-ended questions • Predicting outcomes of daily events and testing their hypotheses • Engaging in trial-and-error learning opportunities • Solving everyday challenges with reasoning and exploration
Social interaction and peer collaboration	• Working together in structured and unstructured activities and learning through peer discussions • Negotiating roles and responsibilities in play and shared problem-solving
Teacher scaffolding and guidance	• Providing verbal prompts and directions to support their learning • Modeling problem-solving strategies • Using engaging methods like songs and interactive communication with children
Free play and creativity	• Expressing imagination through free play, role-play and storytelling • Experimenting with different various materials in an unstructured way

Because qualitative software was not used, I manually reviewed the observation notes, tallied the occurrences of each code across entries after categorization, and grouped related codes into themes based on conceptual similarity and recurrence. The data were then analyzed using thematic analysis, with observation excerpts colour-coded into preliminary codes and organized into overarching themes.

Mathematical Concept Development

Children demonstrated emerging mathematical reasoning through recognizing patterns, counting, and comparing shapes and objects in their environment, as shown in Table 25. They frequently pointed out geometric features "*I see a triangle in the window*" (ON, Charlotte's group, March 18, 2024) or "*These are the same*!" (ON, Darcie's group, March 25, 2024) and used comparative language like "big," "small," or "taller." Teachers occasionally prompted identification, such as "*Can you find a square*?" (ON, Agnes's group, March 26, 2024) leading to responses like bringing a cube labeled as a square. In one structured 4-week control group, children counted sides and vertices of polygons ("*How many sides does the hexagon have*? *Six*!") (ON, Darcie's group, March 22, 2024), though teachers used everyday words like "corners" instead of "vertices." Overall, this theme revealed that early mathematical concept development was often spontaneous and visually driven, with limited formal vocabulary or guidance before the intervention.

Hands-on Exploration With Manipulatives

Across classrooms, children actively engaged with manipulatives through sorting, stacking, and constructing, revealing how tactile play supported spatial understanding, as presented in Table 25. Examples included arranging coloured sticks into shapes, grouping square and triangular prisms ("*Can you group these blocks*?" "*All squares here*!") (ON, Klara's group, March 28, 2024) and creating bridges with magnetic tiles. Teachers sometimes reinforced learning through corrective questioning ("*That is not a square, that one is a rectangle*") (ON, Charlotte's group, March 20, 2024). Such interactions showed that manipulating materials deepened awareness of shape, size, and symmetry while blending symbolic play ("castle," "tower," "tunnel") with mathematical reasoning.

Inquiry-Based Learning and Problem-Solving

Curiosity and experimentation were evident as children tested ideas and predicted outcomes through guided inquiry as presented in Table 25. Teachers encouraged exploration with open questions like "*Which car do you think will go farther*?" (ON, Victoria's group, March 27, 2024) or "*Do you think it will snow tomorrow*?" (ON, Grace's group, March 22, 2024). Children engaged in trial-and-error reasoning, adjusting sand or gravel to build a "chocolate cake," discovering that "*It won't roll, it's too big*" (ON, Jane's group, March 25, 2024) or using chairs to solve everyday problems. These interactions reflected early hypothesis testing, logical thinking, and flexible problem-solving. Collectively, such moments demonstrated how inquiry and experimentation fostered foundational scientific and mathematical reasoning.

Social Interaction and Peer Collaboration

Collaboration was a recurring theme as children negotiated roles, shared materials, and co-constructed ideas as shown in Table 25. During block play, one child said, "*I need a half here*," and another offered, "*You can have this one*" (ON, Charlotte's group, March 18, 2024) while others cooperatively fit puzzle pieces "*Does this one go here*?" "*Yes, that fits*!" (ON, Agnes's group, March 26, 2024). They resolved conflicts ("*You have too many trains!*" "*Okay, I'll give you one*") (ON, Darcie's group, March 20, 2024) and took turns during games or imaginative play ("*I'll be the princess*!" "*You can be the other princess*!") (ON, Jane's group, March 18, 2024). In more structured tasks, peers debated counts and measurements ("*We need five more. ... No, two more!*") (ON, Darcie's group, March 20, 2024). These examples illustrated how collaborative dialogue supported reasoning, fairness, and shared problem-solving.

Teacher Scaffolding and Guidance

Teachers consistently modeled, prompted, and guided children's reasoning through targeted questioning and demonstration as presented in Table 25. For instance, "*What happens if you put two triangles together*?" led to the response, "*A diamond*!" (ON, Darcie's group, March 27, 2024) and "*Count the sides—are they all the same*?" prompted, "*No, these are longer*!" (ON, Darcie's group, March 27, 2024), reinforcing the concept of rectangles. Teachers used modeling ("*Watch me—these are sides*") (ON, Darcie's group, March 20, 2024) and rhythmic songs like "*A circle is round, round, round*" or "*Draw a diamond in the air*" (ON, Grace's group, March 25, 2024) making geometry language playful and memorable. These strategies encouraged precision, engagement, and gradual abstraction.

Free Play and Creativity

Free play allowed children to integrate imagination with emerging mathematical concepts as shown in Table 25. They built "dragon castles" and "tunnels for trucks," pretended to cook with pebbles, and used cylinders as telescopes or train parts. Through trial-and-error experimentation, packing sand to make stronger castles or mixing paint colours ("*Blue and red make purple*!") (ON, Destiny's group, March 19, 2024) children explored cause and effect and material properties. Even unstructured play, such as measuring or comparing objects, reflected spontaneous engagement with geometric and scientific ideas.

The observations showed that the key aspects young children's daily lives in daycare setting mostly include experiences that were grouped under learning mathematical concepts, interact with their friends and their environment, develop problem-solving skills, collaborate with friends, receive support from their teachers, and express their creativity in both pre-planned and structured and unstructured activities and play as shown in Table 25. These themes all together highlighted the various ways in which young children understand mathematical concepts, problem-solving, and social learning and provides a thematic analysis of the general experiences in daycares in Niagara Region. The themes are explained with greater detail and more examples from the observation notes in the following sections. Building on this baseline, the next section presents findings from the second round of observations conducted during the intervention, highlighting how the FG-EGP reshaped children's engagement, reasoning, and collaborative exploration of geometric ideas.

Second Round of Observations (During the Intervention)

During the second round of observations, conducted during the FG-EGP intervention, two core strands of learning became particularly pronounced. The analysis followed Braun and Clarke's (2006) approach, beginning with inductive open coding of observation field notes. Initial codes were assigned to specific observed behaviours and verbal expressions (e.g., "touching sides," "naming 3D shapes," "flipping shapes," "asking for triangles," "negotiating shape names"). These codes were then grouped into subthemes based on conceptual similarity and frequency, forming hierarchical categories within each overarching theme.

Table 26

Emerging Themes and Subthemes from the Second-Round Observations

Themes	Subthemes (codes)
Development of geometric reasoning with manipulative-based exploration and concept building	• Recognizing and naming geometric shapes • Understanding properties and relationships between shapes • Exploring spatial relationships and measurement • Constructing and decomposing shapes with manipulatives • Hands-on learning for abstract concepts
Peer collaboration and communication in mathematics	• Working together to solve spatial and logical challenges • Verbalizing mathematical thinking in group work

Under *Development of Geometric Reasoning*, codes were clustered into five subthemes: Recognizing and Naming Geometric Shapes; Understanding Properties and Relationships Between Shapes; Exploring Spatial Relationships and Measurement; Constructing and Decomposing Shapes with Manipulatives; and Hands-on Learning for Abstract Concepts as presented in Table 26. These subthemes emerged from frequent observations of children identifying, comparing, composing, and transforming shapes using Froebel Gifts and teacher scaffolding. Observational excerpts showed teachers guiding children through side-counting, surface comparison, shape flipping, and component analysis. For example, the prompt "*Touch your sides*" frequently coincided with children correctly identifying sides and vertices, reflecting emergent abstraction. Children actively engaged with physical shapes through probing questions and tactile experiences. Teachers supported their emerging abstraction by encouraging them to count sides, distinguish between 2D and 3D forms, and build new shapes from manipulatives like Froebel Gifts. Dialogue and hands-on guidance (e.g., "*Touch your sides*," "*Flip your shapes*") helped transform visual recognition into an understanding of properties such as edges, faces, and symmetry.

The second major theme, *Peer Collaboration and Communication in Mathematics*, was derived from codes that captured children's social interactions during group tasks, including collaborative building, shape sorting, and shared reasoning as shown in Table 26. Two subthemes emerged (a) Working Together to Solve Spatial and (b) Logical Challenges and Verbalizing Mathematical Thinking in Group Work. These codes appeared in instances where children corrected each other's misidentifications, jointly constructed shapes, or used geometric language in peer dialogue (e.g., "*That's not a square. That's a rectangle*!") (Observation notes (ON), 8-week experimental group, May 8, 2024). Musical reinforcement and shared verbal justification during shape classification tasks also contributed to this theme. Through shared tasks, like constructing bridges or composing a square from triangles, they spontaneously used geometric vocabulary ("*straight sides,*" "*corners*") and corrected one another. These moments of collaborative reasoning not only reinforced individual abstraction but also highlighted the role of social interaction in making geometric thinking more explicit and meaningful. The thematic structure captured both the individual abstraction processes supported through manipulative use and the socially constructed nature of children's geometric understanding within the FG-EGP learning environment. To complement the observational insights, the following subsection examines teachers' perspectives through semi-structured interviews, providing deeper understanding of their instructional decisions, vocabulary use, and manipulative integration across different groups and durations.

Teacher Interviews: Overview and Thematic Focus

As part of the qualitative data collection, semi-structured interviews were conducted with a total of nine teachers across the participating daycare centres. These teachers were grouped based on their role in the study's design: five teachers represented the control group (Victoria, Jane, Darcie, Agnes, and Grace), and four teachers represented the experimental group (Destiny, Charlotte, Gisele, and Klara). The study also accounted for intervention duration; among the control group, four teachers participated in the 4-week phase and one (Grace) in the 8-week phase, while in the experimental group, one teacher (Destiny) was part of the 4-week implementation, and three teachers (Charlotte, Gisele, and Klara) were part of the 8-week implementation. These interviews aimed to uncover teachers' instructional strategies, use of manipulatives, and geometric vocabulary practices before, during, and after the intervention.

Interview transcripts were analyzed thematically to identify patterns across experimental and control groups, to explore how instructional conditions aligned with the development of abstraction in children's shape recognition, and to capture the instructional dimensions that shaped children's engagement with geometric concepts. Three overarching themes emerged from the analysis: (a) *Instructional Factors in Geometry Learning*, (b) *Manipulatives and Their Role in Learning*, and (c) *Vocabulary and Geometric Discourse*.

Table 27

Teachers' Interview Themes

Themes	Subthemes
Instructional factors in geometry learning	• Prompting children to think about abstract properties of shapes • Prompting connections between abstract properties of shapes • Connecting abstract geometry concepts to known content from children's lives • Facilitating shape-comparing activities
Vocabulary and geometric discourse	• Using appropriate vocabulary when describing shapes • Prompting children to use appropriate terminology when describing shapes • Facilitating discussions about the properties of shapes
Manipulatives and their role in learning	• Using appropriate manipulatives to represent geometry concepts • Facilitating shape composition activities • Facilitating shape de-composition activities • Supporting children in using manipulatives to represent abstract concepts

Within each theme, further subthemes were identified to highlight specific pedagogical practices. For instance, teachers varied in how they prompted abstract reasoning about shape properties, structured hands-on comparison and composition activities, connected geometry to children's lived experiences, and introduced precise vocabulary. The differences between control and experimental groups were especially pronounced in the areas of abstraction support, vocabulary precision, and use of Froebel Gifts as structured learning tools.

This thematic organization as presented in Table 27 provides insight into how teaching practices differed not only between control and experimental conditions but also by intervention duration. The 8-week experimental classrooms, in particular, exhibited more intentional and multi-sensory instructional strategies that integrated manipulatives and vocabulary to support abstract geometric understanding. The subthemes identified within each domain illustrate the nuanced ways in which teacher beliefs, resources, and program exposure shaped the quality and depth of children's geometry learning experiences. The next section organizes teachers' and observational data thematically around the specific types of activities used to engage children with geometry concepts before, during, and after the intervention.

Type of Activities Teachers Use for Children's Engagement With Geometry Concepts Before, During, and After the Intervention

Before the Intervention

Before the intervention, across all four daycare centres, teachers integrated various play-based, inquiry-driven, and manipulative-supported activities into their classroom routines, although the level of structure and intentionality varied. One of the frequently recurring codes in the dataset was recognition of patterns in shapes and objects as presented in Table 27. This code emerged from observation notes where children counted objects, identified or matched similar shapes in their environment which was sometimes guided by the teachers and at times, it was not. Children were often heard making statements such as "*I see a triangle in the window*," "*That one is the same as this one*! (shows a trapezoid)" (ON, Charlotte's group, March 18, 2024) or engaging in shape matching activities during play. In one instance, a child pointed to two pentagon tiles on the floor and said, "*Look! These are the same*" (ON, Darcie's group, March 25, 2024). These moments reflected an early form of geometric reasoning as children noticed visual similarities and repeated configurations around them. These experiences were coded as recognition of patterns in shapes and objects.

Children were frequently observed counting out loud during play, such as when stepping on stones in the outdoor area, counting the sides and corners of the shapes or while setting the table with utensils: "*One, two, three, four…*" (ON, Destiny's group, March 19, 2024). In two classrooms, they worked with worms of varying lengths, covering them with Lego bricks and counting how many bricks it took to cover each one, fostering measurement and comparison skills. "T: *How many bricks cover this worm*? C: Five! T: *What about the longer one?* C1: *Eight!* C2: Ten!" (ON, Jane's group, March 21, 2024). Additionally, they used comparative language to describe differences in objects: "*This is taller*." or "*Mine is tiny*" (ON, Victoria's group, March 22, 2024). Descriptive size-related vocabulary like "*big*," "*short*," and "*small*" appeared often during both structured and unstructured activities. For example, while building with blocks, one child remarked, "*We need the big one now*" and another responded, "*No, it's too big. Take a small one*" (ON, Gisele's group, March 21, 2024). The teachers sometimes asked the children to identify the shapes in their environment such as finding squares in classroom or outside. "T: *Can you find a square for me?* C: Brings a box (cube) and says it is a square" (ON, Agnes's group, March 26, 2024).

In another observation, teachers guided the children by asking them what they were doing or what structure they built. For example, in one daycare, this conversation took place: "T: *What did you build?* C: *This is a castle*" "T: *What shape is this though*? C: *It is like a bit like triangle*" "T: *How many sides does a triangle have*? C: *Three* (touches on random spots on the shape when counting three)" (ON, Grace's group, March 27, 2024). In this observation, the teacher did not guide the child to identify the sides of the shape carefully. Only in one classroom, which was a 4-week control group with the children who were 4 years to 4 years and 10 months old (the oldest group in this study), the children consistently engaged in structured activities such as identifying shapes, counting, tracing and identifying sides and vertices, and recognizing differences between two-dimensional and three-dimensional objects. The teacher of this group asked the children to count the number of sides in various polygons and to point out the sides and the vertices of them. "T: *How many sides does the hexagon have*? C: *6*. T: *Can you trace the sides?* C: *It goes like this, and this and this. T: What about the corners? C: 1, 2, 3, 4, 5, 6* (touches to each vertex when counting)" (ON, Darcie's group, March 22, 2024). However, none of the teachers including this group mentioned the word "vertices" or "vertex" and used the word "corner" to identify the vertices, instead.

Hands-on exploration with manipulatives was common before the intervention. Observation notes revealed that children across all groups engaged intensively in hands on exploration with manipulatives. They frequently sorted, stacked, and classified objects, often during cleanup or guided activities, demonstrating early categorization skills. For example, in a 4-week control group, children coloured ice-cream sticks and arranged them to form shapes, while in both control and experimental groups they grouped square and triangular prisms after prompts like "T: *Can you group these blocks?* C: *Puts all square prisms together and separates triangular prisms*" (ON, Klara's group, March 28, 2024). Magnetic play also featured: "T: *Why don't you put them on the whiteboard and make a bridge that animals can walk upon?*" "C: *Begins to sort out the magnetic blocks one by one on the fridge*" (ON, Jane's group, March 25, 2024). In another session, children helped build shapes: "T: *I am making a square. Can you help* *me?*” “C: *The child makes a perfect square by arranging the magnet strips*” (ON, Victoria’s group, March 27, 2024). Block play supported symbolic representation, building “princess castle,” “tower,” “tunnel,” and so on. Sometimes teachers intervened to reinforce math learning with questions like “*what shape is that*?” “*How many sides does it have*?” (ON, Darcie’s group, March 22, 2024) or corrections like “*That is not a square, that one is a rectangle*” (ON, Charlotte’s group, March 20, 2024). At other times, they’d comment, “*Yes, that is an O, and it is also a circle*” (ON, Destiny’s group, March 19, 2024).

During the Intervention

During the intervention, activities were more structured and intentionally designed to foster geometric reasoning through the use of Froebel Gifts, guided comparison, and hands-on exploration. In the observed classrooms, children engaged in numerous activities that supported the recognition and naming of geometric shapes. Teachers frequently prompted them to identify basic and complex shapes through questions and real-world associations. For instance, one teacher asked, “*What is a circle tower called*?” and then clarified, “*A cylinder*” (ON, 4-week experimental group - Destiny - April 22, 2024). In several instances, children were asked to name shapes shown to them: “*What shape is this*?” “*Is this a circle?*”, to which a child responded, “*Yes.*” The teacher followed up with “*What is a circle*?” and the child explained, “*The sides are curvy*” (ON, 8-week experimental group - Gisele - May 7, 2024). Children also learned to distinguish between two-dimensional and three-dimensional shapes. When the teacher asked, “*What was the sphere like*?”, a child replied, “*A ball*” (ON, 8-week experimental group - Klara - May 7, 2024). Similarly, children referred to a rhombus as a “*diamond*” and a semi-circle as a “*summer circle*" (ON, 4-week experimental group - Destiny - May 1, 2024). They also identified a "*triangle,*" a "*square*," and an "*ice cube*" when shown a cube (ON, 4-week experimental group - Destiny - April 29, 2024).

Teachers guided children through counting sides and identifying defining shape features, often using step-by-step recounting and corrective prompts. Across classrooms, I observed a consistent scaffolding pattern in which teachers first invited an estimate, then supported children in re-checking through touch and counting.

- Destiny (4-week experimental group) prompted children to count triangle sides. One child initially overcounted ("*4! 1, 2, 3, 4*!"), but after guided recounting correctly identified three sides. When a child counted five sides for a square by including the front face, Destiny clarified the distinction by explaining, "*No, those are faces*," redirecting attention to edges rather than surfaces (ON, April 25, 2024).
- Klara (8-week experimental group) emphasized property differences between two- and three-dimensional figures. In a comparison of a circle and a sphere, she prompted children to reason about functional properties. A child concluded that the circle was different because "*It is flat. It doesn't roll*," showing emerging attention to defining attributes rather than appearance alone (ON, May 7, 2024).
- Charlotte (8-week experimental group) led structured comparison discussions across sets of triangles. She repeatedly asked whether triangles were the same and what differed, guiding children to attend to measurable variation. Children responded using comparative language such as size differences ("*some of them are small, some of them are bigger*") rather than relying only on visual prototype judgments (ON, May 13, 2024).

Teachers also embedded spatial relationship and measurement language within matching and sorting tasks. I observed prompts that directed children's attention to side length and equality rather than overall appearance. For example, Charlotte guided matching by contrasting "shorter sides" and "equal sides," while Klara used tactile prompts such as asking children to touch edges and decide whether they were "flat and straight" or "bumpy, curvy," thereby linking sensory inspection to geometric properties (ON, May 14–15, 2024). In Destiny's classroom, when comparing a semicircle and a triangle, a child identified the semicircle by reasoning that "*one is half*," correctly associating the term with the curved figure (ON, May 2, 2024).

Composition and de-composition tasks were frequently structured as shape-building challenges. Across sessions, I observed teachers posing combination problems and inviting prediction before assembly:

- Destiny asked what would happen if two semicircles were combined; children concluded they formed "*a circle*."
- When two triangles were joined, children predicted and then confirmed that they made "*a bigger triangle*."
- Open construction challenges ("*Who can create a giant triangle using different triangles?*") led children to produce composite figures they named as "*a diamond*," "*a square*," or familiar objects such as "*a pizza*," sometimes describing part–whole structure (e.g., "*a diamond with 2 squares*") (ON, May 1, 2024).

Narrating these exchanges revealed a consistent instructional pattern across Destiny's, Klara's, and Charlotte's classrooms: teachers moved from estimation to guided checking, property-focused language and composition challenges, systematically shifting children's attention from visual appearance toward structural attributes and part–whole relationships.

From the interviews, teachers described how they supported vocabulary and shape recognition:

Destiny: "*So when they made their creative shapes using the flat pieces, the flat shapes, the colourful ones... they would often try to say those words. We did an activity where we would do a walk around the classroom to find the sphere and the cube. And some of them said 'that is a ball' or 'that is a cube,' and I was saying 'yes, it's a ball, but it's a sphere.' One day, we did a game in circle time where we put up two shapes and talked about which ones were the same or different... and then we talked about why.*"

Gisele: "*I would start with just one shape at a time. Like if we were just exploring the triangle, we would build something just using the triangle shapes. And then the next day we might just use the rectangle or just the square. And then eventually we would combine them to create more complex shapes.*"

Charlotte: "*I asked them what shape they used and what it looks like. They would say, 'I used a square and triangle, and it looks like a rocket.'*"

To complement the quantitative findings, Destiny's detailed reflections on her teaching reveal multiple moments where embodied interaction with geometric materials facilitated abstraction. She stated, "*showing the shapes and then having them touch the shapes, playing with the shapes, exploring them was just another level of understanding for them.*" This aligns with principles of embodied cognition (Gordon & Ramani, 2021; Maturana & Varela, 1987), where conceptual understanding is constructed through bodily experience rather than passive observation. Rather than merely seeing a triangle or a square on paper, children were invited to feel, construct, and manipulate the properties of shapes.

The use of Play-Doh, sticks, and Froebel Gifts became a medium through which abstract ideas like edges, sides, and angles were physically enacted. Destiny noted, "*if I just sat down with some students and said, okay, where's the faces? They'd be like, what do you mean?*"

Instead, students made sense of the concept through active construction: "*let's make it and let's touch it… They want to do it themselves.*" This approach supports Gordon and Ramani's (2021) model of embodied cognition, which highlights how gestures and material manipulation serve as external representations that reduce cognitive load and deepen conceptual understanding.

A particularly illustrative moment came when Destiny explained how children used curved sticks and Play-Doh to explore angles: "*to make a triangle they need to look like a house almost… they couldn't be straight up, otherwise it couldn't make that shape.*" Here, children's understanding of angle and form emerged through the physical positioning of elements, revealing that their learning was grounded in action and perception, not just symbols or language.

Additionally, Destiny introduced the concept of vertices by having children physically touch the pointed parts of shapes and say *"ouch"* to associate sharpness with corners. "*We again, touched our chins. We made the whole 'ouch' concept… they would get the concept like when two sides came to a point it would have vertices.*" This sensory cue anchored vocabulary learning in bodily sensation, a hallmark of embodied learning approaches. These moments demonstrate that the FG-EGP not only facilitated cognitive abstraction through structured, property-focused instruction but also engaged embodied cognition pathways by making abstract geometric ideas tactile, visual, and kinesthetic.

Peer collaboration was embedded in group activities. "C1: *I will make a house, these are my rectangles!*" "C2: *I will make a roof with these (refers to triangle).*" "C1: *We need to make a bridge. The birds and raccoons and the squirrels are going to walk on it*!" "C2: *Let's use the rectangles. They are long!*" "C1: *But they will fall. Put this (refers to a triangle) under them*?" "C1: *Like this*?" "C2: *Yes! It is good*!" "C1: *This is not a pyramid. It looks like a square. Right Miss. Gisele*?" "T: *A pyramid has triangles on the sides. Maybe if we put triangles here*?" "C1: *Yes! Eli, do the same*!" "C1: *I want to make a big square, but I only have these two triangles.*

Beatrice has more." "C2: *Maybe if you put them together*?" "C3: *They fit! But it looks like a rectangle*." "T: *What if we flip them like this*" "C3: *Yes*!" "T: *The two triangles make one big square*!" (ON, 8-week experimental group - Gisele - May 16, 2024).

After the Intervention

Following the intervention, both control and experimental group teachers reflected on their geometry practices and children's engagement. However, marked differences emerged in the complexity, structure, and conceptual depth of the activities they described. While control group teachers described incremental improvements within traditional practices, experimental group teachers, who implemented the FG-EGP using Froebel Gifts, reported more intentional, multi-sensory, and conceptually grounded activities that fostered children's abstract understanding of geometric properties.

Control teachers continued to promote shape recognition through matching games, open-ended play, and real-life object identification. Grace (8-week) described enriched activities involving real-world links, such as reading a book ("*We read a book about Terry Triangle and he has an adventure…*") and using pizza slices ("*when you cut the pizza up…we're cutting it into triangles*") to support geometry connections. She encouraged active discovery, having children search the classroom for shapes: "*We would suggest a shape and the children would have to go around the room and find that shape*." Grace also incorporated sensory and comparative tasks to explore shape features: "*how many sides and… whether it was smooth…*" and "*A ball. It's a circle or hula hoop*."

Other control group teachers, such as Victoria, Jane, Agnes, and Darcie, continued to facilitate shape comparisons and matching through standard activities like mirroring, mat work, and verbal corrections, though these practices generally lacked systematic use of instructional tools or abstraction-focused strategies. Agnes noted that "*the children mostly noticed that the fish* *tank was a rectangle… door… smart board.*" Darcie contrasted shapes such as parallelograms and trapezoids, while Jane clarified the difference between rectangles and squares during comparison activities. Overall, the control classrooms remained grounded in environmental noticing and descriptive labeling. Among them, Grace and Darcie stood out for integrating more structure and comparative reasoning into post-intervention geometry learning.

Teachers in the experimental group, who implemented FG-EGP with Froebel Gifts and structured shape exploration strategies, described more conceptually rich and varied activities. Destiny (4-week experimental group) prompted children to relate abstract shape properties to real-world items and construction tasks: "*what do you have that has sides or these shapes… is it a cube?*" (ON, 4-week experimental group - Destiny - April 29, 2024) She highlighted how children began articulating properties through art: "*when they were creating… making houses… they understood the sides and the shapes and… how they are flat*," and noted how they generalized geometric properties: "*Square… relate it to things like a window… once we taught them about the sides, they would talk about the sides, four sides, flat sides…*"

Charlotte (8-week) emphasized vocabulary-to-structure alignment: "*They said diamond… they knew that that was another word for diamond,*" and supported construction with guided representations: "*I'm gonna make my house… the roof has a corner. Can you make the corner on the roof?*" (ON, 8-week experimental group - Charlotte - May 10, 2024). She embedded comparison and counting into play, such as with puzzles and a shape-counting caterpillar.

Gisele (8-week experimental group) grounded geometric abstraction in daily life and material engagement: "*They make connections with real life objects… 'Oh, the tire has a circle. The ball is a sphere.*" She encouraged property-based reasoning: "*They will compare with items* *that they know… say that it was a diamond… a kite,*" and integrated building materials to deepen conceptual understanding.

Klara (8-week experimental group) engaged children in physically embodied exploration and shape comparison using ramps, movement, and sensory feedback. She reported, "*We would run down the ramps with their own bodies... so they could feel it in their own bodies physically,*" and encouraged both creative building and structured matching: "*Some of our children did build shapes. But most... just built pictures of their own creations.*"

In contrast to the control group, experimental teachers did not rely on incidental shape identification but intentionally embedded geometry into broader cognitive, sensory, and linguistic development. Even though both groups of teachers encouraged shape-related activities after the intervention, control classrooms maintained an emphasis on recognition, informal comparisons, and environmental noticing. In contrast, experimental classrooms, particularly in the 8-week implementation, shifted toward explicit abstraction support, multi-sensory comparisons, and structured manipulative use, reflecting the pedagogical influence of the FG-EGP and Froebel Gifts. These distinctions highlight the potential of intentional, curriculum-based instruction in fostering early abstraction in geometry learning. Since vocabulary plays a vital role in shaping children's conceptual understanding, the following subsection explores how teachers' and children's geometric language evolved throughout the intervention phases.

Vocabulary Teachers Use for Children's Engagement With Geometry Concepts Before, During, and After the Intervention

Before the Intervention

Based on the first-round observations, geometric vocabulary in classrooms was limited in precision and frequency, with most teachers and children relying on informal or basic terms.

Before the implementation of the FG-EGP, control group teachers such as Victoria, Jane, Darcie, Agnes, and Grace demonstrated varying degrees of vocabulary use in geometry instruction. Victoria focused mainly on basic 2D shape names and avoided introducing advanced terms. She stated, *"They use the square and rectangle*" and noted, "*Most of them know the rectangle and triangle but sometimes three of them has like no idea what the rectangle is*." She explained how she described a trapezoid: "*Like most of them I had no idea but once I was asking to gather maybe and say it looks like triangle but the top is cut*." Teachers occasionally prompted vocabulary use, but in most cases, children's shape descriptions were spontaneous and rooted in everyday language. Children commonly used names like "*triangle*," "*square*," or "*circle*" when recognizing shapes in their environment (e.g., "*I see a triangle in the window*") (ON, Charlotte's group, March 18, 2024). They also referred to 3D objects as two-dimensional shapes ("*Brings a box (cube) and says it is a square*") (ON, Agnes's group, March 26, 2024). Teachers often accepted such informal labels without correction or clarification. Children frequently used terms such as "*big*," "*small*," "*short*," and "*tall*" during play, and counting language ("*One, two, three, four…*") emerged when measuring or constructing shapes. Teachers reinforced counting sides or corners occasionally (e.g., "*How many sides does the hexagon have?*") (ON, Darcie's group, March 22, 2024) but typically substituted formal terms like "vertices" with colloquial ones like "*corners*." Only in one structured classroom (4-week control group) did teachers systematically encourage counting sides and corners: "T: *How many sides does the hexagon have?* C: *6.* T: *Can you trace the sides?* C: *It goes like this, and this and this.* T: *What about the corners?* C*: 1, 2, 3, 4, 5, 6*" (ON, Darcie's group, March 22, 2024).

However, even in this setting, the teacher did not use terms like "vertices," defaulting instead to "*corners*." In summary, before the intervention, geometric vocabulary use was mostly limited to informal references and basic shape names, with few instances of structured, explicit vocabulary instruction.

During the Intervention

The second-round observations revealed a noticeable shift in vocabulary richness and accuracy in experimental classrooms during the FG-EGP implementation. Children were observed actively using geometric terms, often introduced and reinforced by their teachers during hands-on activities. Teachers consistently asked questions like "*What shape is this*?" or "*What is a circle?*" to which children replied using descriptive terms like "*The sides are curvy*" (ON, 8-week experimental group - Gisele - May 7, 2024).

Children learned to distinguish between shape types and their properties: "T: *What was the sphere like?*" "C: *A ball.*" "T: *What happens if you combine these two triangles?*" "C: *It makes a big triangle*!" (ON, 4-week experimental group -Destiny- May 1, 2024).

Teachers guided children to correctly count and name properties (e.g., "*No, those are faces,*" when a child mistakenly included the front face of a square in a side count). They scaffolded understanding by prompting clarification: "*Can you show me the long sides?*" "*Now can you touch the short side*" (ON, 8-week experimental group - Klara - May 16, 2024). In peer collaboration contexts, children used terms like "*straight sides*," "*rotate*," "*diamond*," "*square*," and "*triangle*" during creative construction tasks: "C1: *I want to make a big square, but I only have these ones (two triangles).*" "T: *What if we flip them like this?*" "C3: Yes!" "T: *The two triangles make one big square!*" (ON, 8-week experimental group - Gisele - May 16, 2024). During the FG-EGP intervention, experimental group teachers (Destiny, Charlotte, Gisele, Klara) emphasized accurate geometric vocabulary through repetition, guidance, and structured dialogue. Destiny noted that children adopted 3D shape names: "*I think I introduced that vocabulary a couple times, but they would use the actual names of the 3D shapes more likely than saying 3D.*

Like they really got the concept of sphere and cube..." She highlighted how children distinguished between shapes: "*we did emphasize the fact if they're flat and straight, they use the terminology straight a lot or curved...*" To reinforce learning, Destiny explained, "*anytime that I saw a shape that was the one that we were trying to get out of them, I would repeat it a hundred times.*" She described how vocabulary emerged through activity: "*They also like the word smooth... Because when we were doing just the sphere... we discovered that some things can roll and some things can't.*"

Charlotte reported that the children mirrored her terminology: "*If I was using that kind of language, so were they," and "They've always also said corners. Umm. Sides... when they circle, they knew that the circle has no lines, so they said it was round.*" She encouraged refinement, as when she corrected "*diamond*" to "*rhombus*": "*Oh yes, I didn't make this the proper term for a diamond. It is a diamond shape, but the proper name is rhombus*." Charlotte helped children describe features: "*They said it had two flat circles, and then it has one on top, one on the bottom...*" and "*That angle has two sticks and a corner*."

Gisele described how she redirected children's informal descriptions: "*For sphere, they mostly said 'ball.' We did a lot of redirecting and said, 'That's a ball, but it has a sphere shape.'*" She modeled vocabulary use: "*Well, it was redirecting them to use the proper language because they will say it's a square. It's a circle and then we will use the proper words for sphere and cube*." Gisele encouraged reasoning: "*They will mention the sides of the cube, saying how you have to have all the lines to make the cube.*" She guided children from visual references to formal language: "*They just said about the shape that it was just pointy and that it had like the pointy edges and that's it.*"

Klara also modeled and corrected language consistently. She noted: "*We were referring to them as three-dimensional shapes, but the kids were often calling them big shapes.*" She emphasized tactile learning: “*The most that they got out of it was the feel of the texture... where it was flat or pointy…*” Klara guided exploration with prompts like: “*How can I manipulate this object? Or what colour is it? Where's basically where's its family? Where's his family?*” When children confused 2D and 3D terms, she clarified: "*They would often call them by their two-dimensional shapes and we would tell them the three-dimensional shape.*"

These excerpts suggest that during the intervention, children developed a growing ability to use and understand geometric terms with the help of hands-on manipulation, guided questioning, and peer interaction.

After the Intervention

Interview data from teachers in the experimental and control groups (post-intervention) confirmed the patterns observed during the second round and offered insight into how vocabulary use evolved. Most control group teachers reported relying on basic shape names. Agnes shared: “*For the cube, either box or cube. They would just say cube or box.*”
Jane stated: “*The only type of language we would use with them would be like, you know, the shape, it's a square, it's a circle, it's a whatever.”* Darcie, however, modeled precise terminology and corrected children: “*I would correct them ’cause we always use proper words… I would say it would be round or it would be a sphere. Like I would add those words and for them to mirror back.*” Grace alternated between informal and semi-formal terms like “*edges*” and “*corners*,” and acknowledged some 2D/3D vocabulary (e.g., “cube” vs. “cuboid”).

Teachers in the experimental group consistently emphasized correct vocabulary and corrected children’s informal descriptions: Destiny shared: “*I think I introduced that vocabulary a couple times, but they would use the actual names of the 3D shapes more likely than saying 3D. Like they really got the concept of sphere and cube.*” Charlotte reflected: “*If I was using that kind of language, so were they.*” Gisele added: “*We did a lot of redirecting and said, ‘That’s a* *ball, but it has a sphere shape.'*" Klara supported precision by explaining sensory vs. geometric descriptions: "*It might be explained that when you're saying it's pointing or sharp or pokey, that's how it feels. But that's not the name of it.*"

Teachers credited Froebel Gifts with helping children use formal terms like "cube," "sphere," and "cylinder": Charlotte: "*We would use the Gifts… the wooden shapes and we named them. The cube, the sphere, the cylinder.*" Gisele: "*We talked about what each piece was called… 'This is the sphere,' and they'd say, 'ball,' and we'd say, 'sphere.'*" In summary, post-intervention interviews confirm that experimental group teachers made sustained and deliberate efforts to scaffold geometric vocabulary, using consistent modeling, guided corrections, and manipulative support. Control group teachers (except Darcie) maintained more limited and informal vocabulary use.

Although no formal word-count analysis was conducted, both observation notes and teacher interviews revealed clear qualitative shifts in children's geometric vocabulary during the intervention. For example, children who had previously referred to a *rhombus* as a *diamond* occasionally began using the correct term, often prompted by teachers' modeling. Similarly, several children who initially treated *square* and *rectangle* interchangeably began distinguishing them by referring to side length (e.g., noting that rectangles have "long sides"). These linguistic shifts were consistently noted across observation records and later confirmed by teachers in interviews, indicating that children's use of geometric terminology became more precise and conceptually grounded over the course of the intervention, even without formal lexical quantification. Beyond language, the material tools that mediate learning are equally significant; therefore, the next section focuses on how manipulatives, particularly Froebel Gifts, were used to foster children's engagement and abstraction in geometry.

Manipulatives Teachers Use for Children's Engagement with Geometry Concepts Before, During, and After the Intervention

This section draws on first- and second-round classroom observations as well as post-intervention teacher interviews to examine how manipulatives were used in both control and experimental classrooms to engage children with geometry concepts.

Before the Intervention

Before the intervention, while all control teachers incorporated manipulatives in some way, their uses were largely exploratory or supportive of imaginative play rather than intentionally scaffolded for geometric abstraction. For example, Victoria shared that magnetic tiles were always available, stating, "*We just use that… They're obsessed with the magnetic tile*," indicating high engagement but limited instructional intervention. Jane offered two-dimensional shapes and sometimes supported drawing them, saying, "*We would place 'em on the table… maybe do some drawings on the boards of it as well,*" which hints at potential for shape recognition but lacks follow-up with conceptual vocabulary or spatial reasoning. Darcie showed more structured use by regularly rotating tactile materials "*We've done it (angles) with sticks, popsicle sticks, natural sticks, the straws and connectors*." yet even this practice often emphasized variety and sensory exposure rather than systematic shape construction or de-composition. Grace offered a more structured approach among the control group, combining blocks with a magnetic board and explicitly referencing shape features like "edges, sides, and vertices," though mostly in the context of identifying or building familiar objects like houses or spaceships. Overall, while manipulatives were abundant across control classrooms, they were generally used for open-ended construction or recognition tasks without a clear conceptual framework linking them to geometric relationships.

Before the FG-EGP intervention, teachers and children in both experimental and control groups used a variety of classroom manipulatives, though often in unstructured or incidental ways. Teachers provided access to magnetic blocks, foam shapes, coloured sticks, and wooden blocks. In most classrooms, these materials were used primarily in play-based contexts without consistent guidance toward geometric reasoning. "T*: Can you group these blocks?" "C: Puts all square prisms together and separates triangular prisms."* (ON, Klara's group, March 28, 2024). In another group, the observation was as follows: "T: *I am making a square. Can you help me*?" "C: *The child makes a perfect square by arranging the magnet strips* (ON, Victoria's group, March 27, 2024). Children engaged in symbolic play with blocks (e.g., constructing castles or tunnels), but teacher facilitation was minimal. In classrooms with older children, there were more deliberate attempts at linking manipulatives to shape concepts. "T: *How many sides does the hexagon have?*" "C: *6*." "T: *Can you trace the sides?*" "C: *It goes like this…*" (ON, Darcie's group, March 22, 2024). Despite some isolated examples of guided use, most classrooms employed manipulatives in spontaneous or informal ways, lacking systematic integration of geometric abstraction.

During the Intervention

In the experimental classrooms during the implementation of the FG-EGP, manipulatives became key instructional tools designed to scaffold children's understanding of abstract shape properties. Teachers used Froebel Gifts systematically to explore relationships between shapes and engage children in building, composing, and decomposing geometric forms. "*What happens if you combine these two triangles?*" "*It makes a big triangle!*" "*Now we are going to combine our shapes and create new shapes*." "*I made a diamond." "I made a square." "A pizza"* (ON, 4-week experimental group - Destiny - May 1, 2024).

Teachers prompted children to physically explore sides, corners, and curves of the shapes through manipulation. "*Touch your sides.*" "*Is it flat and straight?*" "*Touch the edge. Is it flat?*" "*Is it bumpy, curvy?*" (ON, 8-week experimental group - Klara - May 14, 2024). Children worked in groups using blocks and other manipulatives to solve spatial problems collaboratively. "*Let's use the rectangles. They are long!*" "*But they will fall. Put this (refers to a triangle) under them.*" "*Yes! It is good!*" (ON, 8-week experimental group - Gisele - May 16, 2024). These guided manipulative activities supported both visual and tactile recognition of shape properties and fostered more abstract geometric thinking.

Destiny explained in the interview that Froebel Gifts #8 facilitated learning: "*We got to use the sticks and the Play-Doh together, which is cool... we use the paper to mold things.*" She described how wrapping paper around 3D blocks helped children understand the difference between flat and solid shapes: "*So taking that and showing the difference between a flat paper circle and a sphere so that that concept makes sense to them and it makes sense to me.*" Hanging 3D shapes from Froebel Gift 2 also helped children visualize their properties from all angles: "*...you could see all the sides as well, which really indicated that it's not flat*." Destiny found that using sticks made shape construction clearer and more tangible: "*The stick really made it visually easy... seeing the length of the sides, the angles of the sides, it's quite apparent and a lot easier just using it, using your hands*." The Play-Doh and sticks (Gift 10) helped children grasp the importance of equal side lengths: "*...they have to be the same size so that the sides connect to make one big loop.* " She also guided children in composing shapes: "*We made a rhombus or a diamond using two triangles... then we took the triangles out and we used the sticks...*" and using Play-Doh to create 3D shapes: "*...if you didn't have the right angle they would fall down.*" She emphasized the difference between shapes based on angles and lengths: "*...some of them were* *having a hard time doing the triangle on the side... what was stopping them... the length of the sides were all the same so they couldn't get the angles right.*"

Destiny also described de-composition activities: "*We had the different shapes inside the big boxes... I would hold another one up and show the difference.*" She attempted to help children decompose and recompose rectangles from squares, though the children struggled with accuracy. Children explored shapes by surrounding flat forms with sticks: "*...they would take the flat shapes and... go around it... it goes in this shape or form using the edges.*" To support children in using manipulatives to represent abstract concepts, Destiny reported using body associations like touching her chin to explain "vertices" and emphasized hands-on experiences: "*...they started recognizing that the vertices were pushing through the paper... it was pointy or sharp.*" She believed the children benefited from these experiences even if they lacked vocabulary: "*...they naturally did and figured out on their own but they may just not have known the vocabulary for that.*"

Charlotte used popsicle sticks, pipe cleaners, and light-up whiteboards to support children's understanding. She also used Froebel Gifts 8 and 9, explaining an activity where children placed small beads (Gift 9) at vertices of the shapes: "*...just to see if they would remember where to put the little round circles, where the points of shape was.*" She facilitated composition by having children copy drawn shapes using sticks and other materials: "*They made these lines very well... they used the materials very well to me.*" Though de-composition wasn't a structured focus, she observed informal moments: "*...it looks like a triangle but they cut the top.*" Charlotte guided children in using manipulatives to build 3D shapes: "*...the blocks and the sticks*." She also introduced magnetic tiles: "*...they have shapes, but they're also magnetic... they learn their shapes.*" Gisele incorporated Froebel Gifts 1 to 6 and emphasized building with cubes. She facilitated composition through block play and stick-Play-Doh activities (Gifts 8, 9, 10): "*...that was the best example that we could use...*" She did not describe structured de-composition activities but used Play-Doh and sticks to help children understand that "*...if you don't put one side, then you're not going to have the shape.*" She guided the use of points (Gift 9) and sticks (Gift 8) to explore sides and angles: "*...that was a really critical activity for them to understand... all the sides of the shapes.*"

Klara integrated Froebel Gifts with regular classroom manipulatives both indoors and outdoors, including the mud kitchen. Children explored concepts like weight, movement, and sorting: "*...we set them down ramps so that they could... explore the way that they moved*." Klara observed that 3D shapes, often reserved for older children, were valuable when introduced early: "*...it's not something that was brought into everyday life for them.*" She described how Gift 8 and 2 from Froebel Gifts helped children feel edges and vertices. She emphasized sensory exploration over cognitive explanation: "*...they were more interested in the feeling and the more... manipulating and the building part*." De-composition was not a major focus, but some sorting and tracing activities occurred. Children counted surfaces and categorized by shape and colour. Klara used structured and exploratory activities to guide understanding of abstract concepts, combining Froebel Gifts with loose parts like beads and rulers to explore size and comparison.

After the Intervention

Post-intervention interviews further revealed the contrast in how manipulatives were used in control and experimental classrooms. Several control group teachers indicated that manipulatives were used more for exploration than targeted geometry instruction. Victoria: "*We just use that… They're obsessed with the magnetic tile.*" Agnes added: "*They're able to use a circle and put that into their Play Doh to make a circle shape*." Grace noted: "*We used blocks to talk about edges, sides, and vertices…*" Darcie, more intentionally, used everyday materials for conceptual understanding: "*We've done it (angles) with sticks, popsicle sticks, natural sticks, the straws and connectors…*"

In contrast, experimental teachers emphasized the role of Froebel Gifts and other materials in deepening children's understanding of geometric ideas. Destiny said: "*Having the actual solid shapes, the 3D shapes, and then putting the paper around it… showing the difference between a flat paper circle and a sphere.*" "*The stick really made it visually easy. They could clearly see, oh, I'm missing a side…*" Charlotte added: *"We counted the side, the lines, and then we counted... they had to put the little small points down on the corner."* Gisele described her teaching experience during the FG-EGP as: "*The one with the sticks and Play-Doh was the best example. ... That was really critical for them to understand that concept about all the angles, all the lines, all the sides of the shapes.*" "*Gift 5 was really great for them… more meaningful to them to recognize the shapes.*" Klara: "*They were able to cook with the two-dimensional shapes*." "*They weren't really interested in the cognitive part… more interested in the feeling… the rolling and the manipulating*." These reflections underscore that in the experimental classrooms, manipulatives, especially Froebel Gifts, were not just play tools but foundational components of structured geometry instruction. Integrating these qualitative insights with the quantitative results presented earlier, the next section synthesizes the findings to illustrate how structured, multi-sensory instruction with Froebel Gifts collectively advanced both cognitive and linguistic aspects of geometric abstraction in early childhood.

Integration of the Qualitative and Quantitative Findings

In conclusion, the findings from both quantitative assessments and qualitative observations converge to affirm that structured, conceptually rich, and multi-sensory instruction, particularly when sustained over eight weeks, can substantially enhance young children's geometric shape recognition and mathematical language. The 8-week experimental group demonstrated the most impressive outcomes, with large effect sizes in both shape accuracy and verbal reasoning, supported by immersive classroom environments that featured guided manipulative play, repetitive vocabulary scaffolding, and collaborative problem-solving. In contrast, shorter interventions or standard instruction, while yielding modest shape-learning gains, failed to cultivate significant growth in verbal expression or deeper abstraction. These results highlight not only the importance of intentional teacher practices, such as the strategic use of Froebel Gifts and real-world analogies, but also the critical role of intervention duration in facilitating meaningful conceptual development. The synthesis of data underscores that true geometric learning in early childhood thrives at the intersection of deliberate pedagogy, rich manipulatives, and sufficient time, a powerful framework for nurturing both cognitive and linguistic growth in mathematics.

The summarized data in Table 18e (Appendix B) show that the 8-week experimental group achieved the most substantial improvement, with a large effect size and significant gains in both recognition and language, attributed to sustained, structured, and sensory-rich pedagogy grounded in the FG-EGP. In contrast, the 8-week control and 4-week groups, both experimental and control, showed modest improvements, with outcomes closely tied to the level and consistency of instructional scaffolding and vocabulary development. This table illustrates how deeper engagement with structured, multimodal learning opportunities, especially when supported by Froebel Gifts, guided abstraction, and intentional language use, correlates with more meaningful advances in geometric thinking and verbal expression.

Table 28

Integrated Summary of the Quantitative and Qualitative Findings

Group	Quantitative change (shape recognition)	Quantitative change (verbal response)	Initial observation	During-intervention observations	Teachers' insights	Integrated interpretation
8-wk experimental	+2.48 pts; Cohen's d = 1.214	Post-test rank 32.67 vs 24.02 (U=270.5, p=.046)	Baseline: counting, shape matching, manipulatives, scaffolding, peer collaboration.	Active geometric exploration: naming/decomposing shapes, building with Froebel Gifts, discussing properties, peer reasoning.	Teachers used structured, sensory-rich, concept-focused pedagogy with intentional scaffolding, vocabulary, and manipulatives to nurture abstraction and discourse.	The strong quantitative gains in shape recognition and verbal ability align with deep, intentional instructional practices. The structured multimodal interventions clearly supported significant improvements in both concept understanding and language skill.
8-wk control	+0.43 pts; Cohen's d=0.609; t=−2.53, p=.018	No significant change reported	Basic counting, shape play, teacher prompts, social problem-solving.	Standard activities around shapes, songs, block play, and incidental scaffolding, without FG-EGP-specific strategies.	Teachers used general scaffolded geometry instruction such as counting, naming, use of real-life analogies and partial vocabulary support, but less consistency in abstraction and concept depth.	Quantitative gains were modest and verbal improvement negligible, reflecting the relative lack of intensive, structured geometry instruction and vocabulary reinforcement. Standard pedagogy yielded limited growth.

Group	Quantitative change (shape recognition)	Quantitative change (verbal response)	Initial observation	During-intervention observations	Teachers' insights	Integrated interpretation
4-wk experimental	+0.47 pts; Cohen's d=0.640; t=−2.824, p=.014	Post-test rank lower vs 8-wk; U=125.5, p=.023 vs 8-wk	Same baseline behaviours.	Brief FG-EGP exposure likely offered some exploration but was limited by duration.	Teacher employed intentional pedagogy, multi-sensory tools, vocabulary, Froebel Gifts, but had limited time to deepen abstraction and peer mathematics discourse.	Though improved over control, the brief intervention wasn't sufficient to yield robust gains. Teachers' intentional methods showed promise but were constrained by duration, explaining mid-level quantitative outcomes.
4-wk control	+0.35 pts; Cohen's d=0.587; t=−2.666, p=.015	Verbal initially higher and remained so	Structured, teacher-led geometry from outset: polygons, sides, vertices, 2D vs 3D.	Continued structured lessons with consistent, but not highly abstracted, geometry focus.	Teachers varied: Darcie and Grace used intentional scaffolding and vocabulary; Victoria and Jane focused on basic naming; Agnes was more incidental.	Gains in shape recognition mirrored the structured baseline, but lack of depth and vocabulary challenge limited both shape and verbal advancement. Teacher variation drove uneven outcomes.

This study examined how the FG-EGP, an instructional intervention inspired from Froebel Approach and Gifts, influenced the abstraction processes of 3- to 4-year-old children in early childhood classrooms. Using a concurrent mixed-methods design, the study investigated whether structured geometry activities, centred on composition, de-composition, and comparison, would support a cognitive shift from visual to property-based reasoning. Quantitative results from the Geometric Shape Recognition Test (GSRT) and qualitative data from classroom observations and teacher interviews were integrated to provide a comprehensive understanding of how children's shape recognition and abstraction developed in different instructional contexts.

The quantitative analyses demonstrated that children in the 8-week experimental group showed the most substantial gains in shape recognition scores, with statistically significant improvements from pre- to post-test and a large effect size (Cohen's $d = 1.214$). The 4-week experimental group also showed significant but smaller improvements. The control groups (both 4-week and 8-week) exhibited only minor gains, likely due to developmental maturation or test familiarity rather than instructional effects. These findings highlight the importance of extended exposure to structured geometry instruction for promoting meaningful learning outcomes.

However, while overall shape recognition improved, a deeper look into abstraction, defined as recognizing geometric shapes based on their defining properties rather than visual prototypes yielded more nuanced results. Analyses of item-level performance across triangle, rectangle, and square recognition tasks showed that the 8-week experimental group outperformed others in recognizing atypical or skewed shapes, particularly those with unusual orientation or proportions. Though these differences were not always statistically significant, they pointed to a trend toward improved flexibility in children's geometric reasoning. For example, the 8-week experimental group made the highest gains in recognizing skewed and kurtotic rectangles (F = 2.98, p = .036) and approached significance in size-based square recognition (F = 2.53, p = .063).

To examine abstraction more directly, children's verbal responses during shape recognition tasks were coded as "visual," "property-based," or "I don't know." Mann-Whitney U tests revealed that the 8-week experimental group scored significantly higher than the 8-week control group in total verbal responses (Table 61, p = .046; Appendix B), indicating improved verbal engagement with shape reasoning. However, when isolating property-based responses, those in which children explicitly referred to geometric attributes such as "*three sides*" or "*four corners*" no statistically significant differences were observed across groups (ANOVA, F = 1.01, p = .392). This suggests that although children may have become more verbally expressive, they did not necessarily increase their use of precise geometric vocabulary or concepts in a statistically robust way.

Instead, a key finding emerged in the form of gesture-based responses. These embodied indicators, such as children tracing sides with their fingers or pointing to corners while silently counting were classified under visual reasoning but provided evidence of emergent abstraction. A Kruskal-Wallis H test in Table 23 revealed a statistically significant difference in gesture response changes across groups ($\chi^2(3) = 52.05$, $p < .001$), with the 8-week experimental group showing the largest gains. Post hoc Mann-Whitney U tests confirmed that the experimental groups (particularly the 8-week group) demonstrated significantly greater increases in gesture use than their control counterparts. These embodied responses suggest that abstraction was taking root not only in verbal language but also through children's physical interactions with shapes.

Qualitative data reinforced and deepened this interpretation. Observations from the intervention classrooms revealed that children increasingly used manipulatives (e.g., Froebel Gifts, Play-Doh, beads and sticks) to construct, decompose, and describe geometric shapes. They verbalized mathematical thinking in group work, used precise vocabulary such as "*vertices*" and "*edges*," and collaboratively solved design-based spatial problems. Teachers scaffolded this learning by prompting shape comparisons, correcting misconceptions, and encouraging children to use appropriate terms. Children's shift from intuitive, visual naming to more intentional reasoning and explanation was documented not only in their spoken language but also in their embodied actions.

For example, children who previously referred to a square as a "*diamond*" began to identify it correctly when asked to count sides or touch corners. Others discovered that two triangles could form a square or that spheres rolled while cubes did not, articulating these ideas through touch, movement, and group discussion. While some teachers noted that terms like "vertex" were still difficult for children to fully grasp, the frequent use of gestures and hands-on manipulation helped children internalize abstract properties even if they could not yet fully verbalize them.

Thus, the integration of findings in Table 28 suggests that the FG-EGP promoted early abstraction in shape recognition primarily through embodied and manipulative-based engagement, with modest gains in explicit verbalization of geometric properties. Extended duration (8 weeks) played a critical role in consolidating these cognitive gains. The findings also underscore the importance of multimodal instruction in early mathematics education: while verbal abstraction may take time to develop, physical interaction with geometric concepts offers a powerful pathway toward deeper conceptual understanding. The synergy between quantitative and qualitative data in this study affirms the potential of structured, Froebelian-inspired geometry instruction to lay the groundwork for abstract thinking in early childhood classrooms.

The findings presented in Chapter 4 provide converging quantitative and qualitative evidence that structured, manipulative-based geometry instruction through the FG-EGP supports young children's developing abstraction in shape recognition. Statistical analyses demonstrated clear gains in recognition and gesture-based reasoning, particularly under longer intervention duration, while classroom observations and teacher interviews illuminated the instructional mechanisms through which these gains emerged.

Building on these results, Chapter 5 moves from reporting outcomes to interpreting their meaning. This chapter situates the findings within the study's theoretical framework on abstraction, embodied cognition, and early geometric reasoning, and examines how the observed learning patterns relate to prior research. The chapter also discusses pedagogical implications, methodological considerations, study limitations, and recommendations for future research and practice.

CHAPTER FIVE: DISCUSSION

Overview of the Chapter

This chapter presents an in-depth discussion of the study's findings regarding the impact of the FG-EGP on the abstraction processes of 3- to 4-year-old children in recognizing and categorizing geometric shapes. Both quantitative and qualitative results are integrated in the discussion to provide a broader and holistic interpretation of the findings. The discussion continues by connecting the findings to the theoretical frameworks of Piaget, Dienes, and Fischbein and by contextualizing the results within existing research on early childhood geometry learning. In this chapter, I also address the practical implications of the findings for educational practice, curriculum development, instructional material, and teacher professional development. Having outlined the purpose and structure of this chapter, the discussion now turns to the first research question, which explores how differences in instructional practices, materials, and language shaped children's development of abstract thinking in geometry.

Discussion on Results of Research Question 1

Research Question 1 asked: "How did the differences in the geometry activities, vocabulary, and manipulatives used by the teachers impact 3- to 4-year-old children's abstract thinking skills in shape recognition?"

The question serves as the overarching inquiry that frames the study. It explores how variations in instructional practices in the experimental and control groups, particularly those introduced through the FG-EGP, influence children's ability to abstract geometric properties during shape recognition tasks. This broad question is examined through three interrelated sub-questions, each focusing on a specific aspect of abstraction.

The first sub-question asked: “Does FG-EGP have an impact on 3- to 4-year-old children’s abstraction processes in geometric shape recognition? If so, in what ways does it influence their abstraction?” It investigates whether FG-EGP has a measurable impact on children’s abstraction processes and, if so, in what ways this influence is observed. This sets the foundation for understanding the program’s general effectiveness.

The second sub-question asked: “Does FG-EGP affect the frequency of children’s property-based responses in shape recognition tasks? If so, does it lead to an increase or decrease in these responses?” The question narrows the focus by exploring whether FG-EGP affects the frequency of children’s property-based responses, which serve as indicators of their abstract reasoning.

Finally, the third sub-question asked: “Does FG-EGP influence how children use attributes and properties to define shapes? If so, is this influence positive or negative?” The question examines the quality and nature of these responses by asking how children use attributes and properties to define shapes, and whether this influence is positive or negative. Together, the sub-questions provide a layered analysis of the overarching question of progressing from general impact to specific cognitive behaviours and offer a comprehensive view of how structured geometry instruction can support the development of abstract thinking in early childhood.

In general, the findings of this study suggest that the FG-EGP does have a meaningful impact on the abstraction processes of 3- to 4-year-old children in geometric shape recognition. Quantitative results presented in Chapter 4, and the tables including statistically significant gains in shape recognition scores for the 8-week experimental group, indicate that extended engagement with the intervention led to a deeper understanding of geometric forms. These gains were not observed in the control groups or in the shorter 4-week experimental group, highlighting the importance of duration in supporting shape recognition skills. To understand how the FG-EGP influenced abstraction, it is first necessary to examine the overall trends in children's shape recognition performance across experimental and control groups.

Shape Recognition Performance Analysis: Pretest–Post-test Comparisons

Across all four groups, statistically significant improvements in shape recognition scores were observed from pretest to post-test. However, the magnitude and nature of these gains varied across the different group conditions, shedding light on the effectiveness of the FG-EGP intervention and the influence of intervention duration. The 8-week experimental group experienced the most substantial learning gains, highlighting the importance of sustained instructional engagement. In contrast, the control groups and shorter-duration intervention group showed only modest improvements, likely attributable to natural developmental progression or procedural familiarity. These results support the value of structured, discourse and manipulative-based geometry instruction in early childhood settings, particularly when delivered over an extended period with fidelity to the FG-EGP framework.

Similar to my study, Gecu-Parmaksiz and Delialioglu (2019) examined the impact of virtual and physical manipulatives on preschool children's geometric shape recognition. Both modalities improved children's ability to recognize typical shapes such as squares and rectangles, with virtual manipulatives having a slightly stronger effect. However, the study emphasized that children continued to struggle with recognizing atypical shapes (e.g., rotated or skewed triangles), indicating that manipulatives alone are insufficient for promoting deep conceptual understanding. This supports the following findings, particularly the effectiveness of the structured, property-focused, and comparison-rich activities in the 8-week group. It also validates the idea that exposure alone, whether to virtual or physical materials is not enough without guided abstraction and instructional design. The discourse between the teacher and the children plays a very important role here because it facilitates the alignment of children's perceptual experiences with formal geometric concepts through language, questioning, and guided comparison. Vinner's (1993) theory of figural concepts posits that a mathematical concept is understood not only through its formal definition but also through the mental images and intuitive representations a learner associates with it. The discourse between teacher and child helps refine these conceptual images by bridging the gap between informal, perceptual impressions of shapes and their formal properties. Through guided questioning and comparison, teachers help children revise or expand their mental images, shifting from relying on visual prototypes to constructing more accurate, property-based representations of geometric concepts. This refinement process is essential for developing a coherent and flexible understanding of shape categories.

According to Vygotsky's (1978) sociocultural theory, learning occurs most effectively within the Zone of Proximal Development (ZPD), where children can perform tasks with the guidance of a more knowledgeable other. In the context of geometry learning, teachers scaffold children's understanding by using precise vocabulary (e.g., "sides," "vertices," "angles") and prompting children to articulate and refine their own thinking. This verbal interaction supports the development of what Mitchelmore and White (2007) describe as "relational abstraction" where children recognize and generalize similarities across shape examples. Research by Christie and Gentner (2010) further reinforces this by showing that structural alignment through language and guided comparison leads children to focus on relational features rather than superficial ones. Additionally, Verdine et al. (2016) emphasize that labeling and linguistic scaffolding are critical in helping children shift from perceptual to conceptual reasoning. Without such discourse, children tend to rely on static visual prototypes, often misclassifying shapes with atypical orientations or proportions. Therefore, the structured conversations embedded in the 8-week FG-EGP intervention likely played a pivotal role in helping children abstract defining geometric properties, making the learning process both meaningful and cognitively transformative. While these results highlight the effectiveness of the intervention in improving shape recognition, the next section explores the deeper cognitive processes involved, specifically, how the FG-EGP fostered abstraction in geometric reasoning.

Evaluating the Influence of the Froebel Gifts Early Geometry Program on Abstraction in Geometric Shape Recognition Among 3- to 4-Year-Olds

The quantitative findings of this study suggest that the FG-EGP supported more than just children's ability to name or identify shapes. The observed improvements indicate a deeper shift in how children perceive and process geometric forms. According to Fischbein (1993), geometric reasoning involves a dynamic interplay between conceptual understanding (e.g., knowledge that a square has four equal sides and four right angles) and figural representation (e.g., the image of a square always standing on its base). The FG-EGP appears to have fostered children's ability to coordinate these two aspects of geometric reasoning.

Rather than relying exclusively on static prototypes, some children exposed to the FG-EGP began demonstrating what Walcott et al. (2009) describe as "dynamic figural concepts"—flexible mental images that accommodate variations in orientation, size, or proportions. This shift suggests a refining of children's concept images through hands-on experiences, guided manipulation of materials, structured reflection, and exposure to both typical and atypical exemplars. The assessment items used in this study were intentionally designed to capture this shift, incorporating 12 items per shape category (triangle, rectangle, square), including typical, atypical, and non-examples. I began each category with a guiding question such as, “Is this a triangle?” to prompt reflective identification. The purpose of this approach was to elicit children’s underlying concept images and observe how they justified shape membership, revealing the depth of their abstraction rather than mere visual recognition.

Triangle Recognition

My study’s findings align in meaningful ways with the results of Yeşil Dağlı and Halat (2016), particularly in revealing young children’s reliance on prototypical triangle representations and their difficulties in recognizing atypical forms. In both studies, children demonstrated high accuracy when identifying typical triangles, most notably the equilateral triangle in standard orientation, but struggled when triangles deviated from that prototype through changes in orientation, skewness, or kurtosis. This consistent pattern reinforces the idea that early geometric understanding is heavily prototype-dependent and that atypical triangle recognition presents a cognitive challenge for preschool-aged learners.

Yeşil Dağlı and Halat’s (2016) study reported that approximately half of the participating 5- to 6-year-olds failed to recognize triangles with atypical features, with the greatest difficulties arising from rotated or flipped orientations. Similarly, in my study, only about half of the children (≈ 46 of 93) correctly identified atypical triangles. Following the intervention, recognition increased primarily in the experimental classrooms. The 4-week experimental group showed the greatest improvement (skewness gain = +0.20, kurtosis gain = +0.13; Table 14), with approximately three additional children recognizing atypical triangles by the post-test, whereas the control groups displayed negligible or negative change. The orientation gains were minimal across all groups, including the experimental ones, suggesting that rotation-based abstraction remains a persistent challenge even when children are provided with enriched instructional contexts. This finding echoes the interpretation of Yeşil Dağlı and Halat in which they say children are often at van Hiele's "Visualization" level, where recognition is based on the overall appearance of a shape rather than its defining attributes.

Nonetheless, FG-EGP intervention shows promise in extending children's geometric understanding beyond surface-level features. Notably, the 8-week experimental group in my study outperformed others on triangle items involving skewness and kurtosis. The fact that the 4-week experimental group also showed positive gains, especially a +0.2 gain in skewness suggests that even shorter exposure to structured comparison and manipulation can begin to shift children's attention toward the defining properties of triangles. This trend was not observed in the control groups and supports the notion that targeted instruction fosters movement toward more generalized and flexible shape recognition.

In contrast to Yeşil Dağlı and Halat's (2016) participants, my study's experimental group children engaged with Froebel Gifts and structured shape discussions, possibly accounting for the observed gains in recognizing triangles with less familiar attributes. While neither study found statistical significance in atypical triangle recognition, the directional consistency in my descriptive data strengthens the argument that sustained, property-focused exposure and discourse is essential in helping children abstract geometric concepts beyond prototypical constraints. A relevant classroom moment illustrates this shift: "T: (Puts 4 different triangle types, equilateral, isosceles, and right angled and obtuse) "*Now, look at these triangles. Are they the same?*" C: "*No*." T: "*What is different?*" C: "*They are in different colours*." T: "*What else is different?*" C: "*These are small, these are big*." T: "*Look at the sides of this triangle. Some sides are long, some sides are short. Can you show me the long sides?*" Children touch their long sides” (ON, 8-week experimental group - Klara - May 16, 2024). This exchange demonstrates how teacher-led discourse guided children beyond surface-level attributes like colour and size toward identifying defining geometric properties such as side length. Together, these findings reinforce calls from the literature (e.g., Clements et al., 1999; Mitchelmore & White, 2007) to integrate atypical examples, hands-on manipulation, and guided reflection into early geometry education to disrupt prototype fixation and support the development of abstract geometric reasoning.

Rectangle Recognition

These findings suggest that the FG-EGP, especially in the 8-week format, may have enhanced children’s ability to recognize rectangles based on abstract properties rather than visual familiarity alone. Recent research underscores the challenges young children face in recognizing atypical rectangles, those that deviate from the standard prototype through variations in orientation, skewness, or aspect ratio. In a study by Aslan and Arnas (2007b), preschool children aged 3 to 6 were assessed on their ability to recognize basic geometric shapes, including rectangles. The results indicated that children often relied on non-defining attributes, such as the typical orientation or proportions of shapes, leading to difficulties in identifying atypical rectangles. For instance, rectangles presented in non-standard orientations or with unusual aspect ratios were frequently misidentified. This reliance on prototypical representations suggests that without targeted instruction, children may struggle to generalize the concept of a rectangle beyond its most familiar form.

My study’s findings resonate with these observations. While no statistically significant differences were found in recognizing unusually oriented rectangles, descriptive data indicated that experimental groups exposed to the FG-EGP showed higher mean gains in identifying rectangles with skewed and kurtotic variations. This suggests that structured exposure to a variety of rectangle forms and having a conversation about the properties of the shape may support the development of more generalized and property-based recognition skills. For example, one child in the 8-week experimental group placed a rectangular toy car on a square board and stated, "*they are not long like the same, I need a rectangle garage*" demonstrating the use of defining attributes like side length rather than surface similarity (ON, 8-week experimental group - Klara - May 17, 2024). In another interaction, the teacher prompts children to evaluate the shape's features by asking, "*Is it flat and straight*?" and then guiding them to touch and assess the edge directly. When prompted to distinguish whether the edge was "*bumpy, curvy, or straight*," a child responded confidently, "*Straight*" (ON, 8-week experimental group - Klara - May 14, 2024). This moment illustrates how teacher discourse can redirect children's focus from surface-level perception to geometric properties, promoting abstraction.

Clements et al. (1999) explored young children's concepts of shape and found that children often form initial shape schemas based on visual features. Their study highlighted that while children could recognize standard shapes, they struggled with atypical versions, emphasizing the need for educational interventions that focus on the defining properties of shapes. This supports the approach taken in my study, where the FG-EGP provided children with diverse examples and encouraged discussions about shape attributes, thereby promoting a deeper understanding of geometric concepts.

In summary, both my research and existing studies highlight the importance of moving beyond prototypical examples in early geometry education. By incorporating a variety of shape representations and focusing on defining attributes, educators can help children develop a more robust and flexible understanding of geometric concepts, enabling them to recognize shapes regardless of orientation, skewness, or proportion.

Square Recognition

While Halat and Yeşil Dağlı (2016) emphasized that children typically default to visual prototypes, FG-EGP helped push toward more flexible concept images, as seen in the modest improvements in atypical orientation recognition in both experimental groups. Thus, while not statistically significant, my directionally positive results, especially in size and orientation gains, indicate that guided comparison and hands-on engagement began to address the visual rigidity in children's figural concepts that Halat and Yeşil Dağlı (2016) documented. In summary, my results complement and extend Halat and Yeşil Dağlı's findings by showing that a structured intervention like FG-EGP can begin to reshape young children's figural concepts, even with the deeply ingrained prototype bias, by offering sustained exposure to varied square representations.

Interpretation and Implications for Abstraction

These findings support the overarching claim that instructional differences, in particular the structured activities, vocabulary, and manipulative use of the FG-EGP, contributed to children's enhanced shape recognition and emerging abstraction skills. However, it is essential to distinguish between recognition and abstraction. While recognition involves the categorization of shapes, abstraction involves discerning and mentally representing the defining properties that transcend visual appearances. The evidence presented here suggests that the FG-EGP began to support abstraction by moving some children beyond recognition of prototypical shapes to the identification of geometric properties in atypical examples. This was most clearly observed in the 8-week experimental group, which consistently demonstrated higher gains in recognizing skewed, unusually oriented, or disproportionate shapes across all three categories.

Nevertheless, these quantitative results only tell part of the story. They confirm that a shift occurred but do not fully reveal *how* children reasoned through these tasks. Did children begin to rely on properties such as the number of sides, angles, and symmetry instead of mere visual templates? Did they verbalize their reasoning or engage with the manipulatives in conceptually rich ways? To answer these questions and deepen our understanding of children's abstraction processes, it is necessary to turn to the qualitative data; namely, children's verbal responses, classroom observations, and teacher interviews. These data sources illuminate the pathways through which abstraction emerged, shedding light on how children explained their categorizations and how the FG-EGP shaped their reasoning. Although the findings confirm that the FG-EGP enhanced recognition and conceptual understanding, they raise further questions about the ways children expressed their reasoning. The following section examines how often and in what forms children articulated property-based responses.

Influence of the Froebel Gifts Early Geometry Program on Preschool Children's Use of Properties in Shape Recognition

Although the FG-EGP was designed to foster abstract geometric reasoning in preschool-aged children, the overall quantitative results suggest that its effect on increasing the frequency of property-based responses was not statistically significant across the full sample. This finding, while perhaps unexpected, echoes broader challenges in early mathematics education where the development of abstraction is a gradual and complex process (Sarama & Clements, 2009; Sinclair & Bruce, 2015). Young children often require sustained exposure and scaffolded experiences with geometric properties before they can consistently verbalize those properties in reasoning tasks. Therefore, the limited effect across the full cohort may point less to the ineffectiveness of the FG-EGP itself and more to the duration or depth of exposure necessary to transform internalized understanding into verbal articulation.

Despite the lack of significant change across the entire sample, deeper analysis of group comparisons revealed that children in the 8-week experimental group significantly outperformed those in the 8-week control group on verbal response scores. This result is aligned with the literature that emphasizes the importance of sustained, guided instruction using rich, manipulative-based geometry tasks (Clements & Sarama, 2011; van Hiele, 1986; as cited in Teppo, 1991). The prolonged engagement with Froebel Gifts, as integrated within the FG-EGP, appears to have supported children's ability to reason about shape more abstractly. This supports Mitchelmore and White's (2000) argument that the abstraction of geometric concepts emerges more robustly through varied experiences that make structural properties explicit over time.

In contrast, the 4-week experimental group did not achieve similar gains and in fact performed significantly lower than their control counterparts. This disparity underscores the influence of both instructional duration and contextual factors, such as teacher emphasis on geometry and the developmental readiness of children. The 4-week control group included older children and teachers with a documented focus on mathematical language, which may have contributed to their higher performance. This finding is consistent with Dienes's theory of multiple embodiment, which stresses that abstraction emerges when learners experience mathematical ideas through diverse and repeated contexts, a condition more readily met in longer interventions.

Intriguingly, while some children's verbal justifications remained visual in the intervention group, the emergence of gesture-based explanations in the post-test phase suggests that children may have been developing internalized understandings that were not yet fully verbalized. This aligns with the action-cognition transduction (ACT) model proposed by Abrahamson and Sánchez-García (2016), which posits that embodied actions such as pointing, tracing, or manipulating objects often precede and support the development of formal verbal reasoning. As such, the lack of verbal property references in some cases should not be interpreted as a lack of conceptual understanding but rather a transitional stage in cognitive development.

These results reinforce the importance of carefully designed and sufficiently long geometry interventions in early childhood, particularly those that bridge physical manipulation and verbal abstraction. Programs like the FG-EGP that integrate Froebelian materials with guided discourse offer promise, but their effectiveness depends on implementation fidelity, the age and developmental level of children, and the time afforded for concept formation. Future studies might explore how gesture and informal reasoning can be leveraged as stepping-stones toward full abstraction, as well as how teacher training can support the interpretation and scaffolding of such transitional behaviours.

Ultimately, while the statistical findings were mixed, the differences observed between short and long interventions highlight a key implication: fostering abstraction in early childhood is not merely about introducing vocabulary or activities but about cultivating a sustained, reflective environment where children are invited to explore, represent, and eventually articulate the properties that define geometric forms. Beyond verbal reasoning, children also conveyed their understanding through gestures and embodied actions. The next subsection analyzes how these nonverbal forms of expression reveal emerging abstraction and the embodied nature of geometric cognition.

The Influence of the Froebel Gifts Early Geometry Program on Children’s Use of Attributes and Properties to Define Shapes

The findings from this study offer compelling evidence that the FG-EGP supported children’s use of gesture-based responses to communicate geometric properties, pointing to the growing role of embodied cognition (Gordon & Ramani, 2021; Maturana & Varela, 1987) in early mathematical development. In recent years, researchers have emphasized that children’s mathematical reasoning is not confined to verbal or symbolic forms, but often emerges through physical actions that reflect developing concepts (Alibali & Nathan, 2012; Goldin-Meadow, 2011). The marked increase in gesture responses, particularly among children in the 8-week experimental group, suggests that FG-EGP successfully created opportunities for children to engage with geometric properties in ways that transcend traditional verbal instruction.

Gesture, as a semiotic tool, allows children to externalize their thinking before they possess the language to fully articulate it (Church & Goldin-Meadow, 1986). In this study, children traced shapes in the air, pointed to vertices, and mimed angles or curved edges, actions that align closely with the concept of operational knowledge as proposed by Piaget (1969/2006). Rather than defining a triangle as “a shape with three sides,” children who gestured toward each corner or mimicked a triangle's outline were demonstrating emerging internalized understandings of what makes a shape belong to a specific category. These embodied acts often precede and facilitate verbal generalization, reinforcing the idea that gesture is not a byproduct of cognition but an integral part of it.

This interpretation finds strong support in the action-cognition transduction (ACT) model (Abrahamson & Sánchez-García, 2016), which posits that mathematical concepts often evolve through sensorimotor engagement before being codified in language. In the context of FG-EGP, children manipulated Froebel Gifts and used bodily movements to explore shapes, which allowed them to “feel” the properties such as number of sides, angles, or symmetry before naming or defining them. These physical experiences form what Dienes (1971; as cited in Gningue, 2006) would describe as the “embodiment stage” of abstraction, where learners transition from concrete handling to symbolic representation through pattern recognition and transformation.

Moreover, the progression observed in gesture use from the 4-week to the 8-week experimental group underscores the importance of time and repetition in embodied learning. This pattern aligns closely with research on embodied cognition, particularly the model proposed by Gordon and Ramani (2021), which integrates embodied cognition with information processing theory to explain the role of gesture in children’s mathematical reasoning. Their theoretical synthesis highlights that gestures are not merely communicative tools but act as cognitive supports that bridge physical experience with abstract thought. In learning environments, gestures help reduce cognitive load, reveal implicit knowledge, and support generalization to new problems. The substantial gesture gains seen in the 8-week FG-EGP group may reflect this cognitive function, where structured, hands-on activities with Froebel Gifts prompted children to externalize their reasoning through gesture. Gordon and Ramani (2021) emphasize that such embodied actions can aid abstraction by making otherwise invisible cognitive processes visible and accessible. As such, the increased gesture use observed in my study likely signals deeper conceptual engagement and supports the broader literature advocating for gesture-rich, embodied instruction in early mathematics education.

As Mitchelmore and White (2000; 2007) argue, meaningful abstraction is grounded in repeated experiences of structurally similar situations. The extended exposure provided by the 8-

week program appears to have supported deeper integration of bodily actions with cognitive reflection, allowing children to consolidate their operational knowledge and express it in increasingly abstract ways. This finding also aligns with the work of Nemirovsky and Ferrara (2009), who view children's bodily enactments not as peripheral but as essential to the emergence of mathematical ideas.

Importantly, the increased gesture-based responses suggest a positive influence of the FG-EGP on children's readiness to define shapes not merely by appearance but by structural properties. While these gestures may not always have been accompanied by precise verbal language, they represent a developmental bridge between intuitive recognition and formal abstraction. Encouragingly, this embodied reasoning did not displace verbal abstraction but rather appeared to complement it, revealing multiple modes through which children constructed and expressed geometric understanding.

In sum, the evidence points to the FG-EGP's effectiveness in supporting not only verbal abstraction but also embodied, gesture-based reasoning as integral to early geometry learning. By positioning the body as a cognitive resource, the program allowed children to engage with geometry in a developmentally appropriate and meaningful way. These findings underscore the importance of integrating manipulative-rich, movement-oriented instruction into early mathematics education, providing children with constant feedback, where gesture and physical interaction with materials serve as powerful tools for developing abstract thought. While these findings illuminate the children's cognitive and embodied development, it is equally important to understand the instructional conditions that enabled such learning. The following section turns to Research Question 2, which focuses on the teaching practices and materials that shaped children's engagement with geometry concepts.

Discussion on Results of Research Question 2

Research Question 2 asked: “What kind of activities, vocabulary and manipulatives did the teachers use in their geometry activities before and during the intervention?”

Understanding how children engage with geometry before and during a structured intervention offers crucial insight into the development of early spatial reasoning and abstraction. This section responds to Research Question 2 by examining the types of activities, vocabulary, and manipulatives used by educators across both phases of the study, prior to and during the implementation of the FG-EGP. Drawing from classroom observations and teacher interviews, this analysis highlights a distinct pedagogical shift: from incidental, play-based engagement with geometry concepts before the intervention to systematic, cognitively rich, and manipulative-driven learning during the intervention. Before the FG-EGP, children’s learning was shaped by informal exploration with materials like blocks, puzzles, and magnetic tiles, supported by spontaneous teacher questioning and everyday language. During the intervention, however, educators employed Froebel Gifts and intentionally designed tasks to foster deeper geometric reasoning. Children were not only recognizing and naming shapes but were also constructing, decomposing, comparing, and verbalizing their properties through guided play and collaborative inquiry. These shifts illustrate how program structure, instructional language, and material use collectively influence children's geometric thinking and their ability to abstract shape properties beyond visual familiarity. The first instructional barrier addressed through the FG-EGP concerned children’s limited exposure to atypical shapes. The next subsection discusses how the program expanded this exposure and improved recognition through varied examples.

Addressing Barrier 1: Enhancing Shape Recognition through Varied and Atypical Examples

The quantitative results demonstrate that children in the 8-week experimental group showed greater recognition of skewed and kurtotic rectangles and non-standard square orientations, even when the gains were not always statistically significant. This aligns with previous studies (e.g., Aslan & Arnas, 2007b, 2010; Gecu-Parmaksiz & Delialioglu, 2019), confirming that children struggle with atypical shapes without targeted exposure. The FG-EGP directly addressed this barrier by including a rich variety of shape types (typical, atypical, and non-examples) during structured learning sessions. This aligns with my purpose to test the impact of Froebel Gifts on overcoming limited shape exposure. The FG-EGP directly addressed one of the central barriers identified in the problem statement: young children's limited exposure to atypical shapes and non-examples. By integrating Froebel Gifts into daily activities, the program offered children sustained opportunities to explore a wide range of shape types in both structured and open-ended contexts. This was evident in Destiny's classroom, where children engaged with unfamiliar and diverse shapes through hands-on activities such as composing sun shapes with triangles or constructing rectangles from squares. Destiny explained that the arrival of the Froebel Gifts sparked heightened interest: "*Having these boxes come in and being able to show them like their gifts, the kids were so much more intrigued… it was easier for me to introduce the different activities*." Similarly, Gisele emphasized how structured comparisons between spheres and real-world objects like balloons or clocks helped children distinguish between shape and object categories: "*You provided Gift Number One… we compared spheres with balls, balloons, and clocks. … Those connections with real-life items were very helpful.*" Both teachers also supported children's use of precise vocabulary to differentiate between typical and atypical shapes, for example, clarifying the difference between a cube and a square or a ball and a sphere.

These instructional strategies reflect the program's emphasis on providing rich and varied examples of geometric forms, thereby offering children meaningful exposure that supports the development of more inclusive and accurate shape categories. In doing so, the FG-EGP aligns closely with research recommendations calling for variety-rich geometry content and structured shape comparisons to facilitate early abstraction (Christie & Gentner, 2010; Clements et al., 2018; Gecu-Parmaksiz & Delialioglu, 2019). Building on the importance of varied examples, the next subsection examines how structured discussions and teacher questioning further deepened children's reasoning about geometric features.

Addressing Barrier 2: Promoting Geometric Reasoning through Structured Discussion of Shape Variants

While the increase in property-based verbal responses was not statistically significant, children in the experimental groups, especially the 8-week group demonstrated a higher frequency of total verbal engagement ($p = .046$) and a notable increase in embodied gestures when reasoning about shapes. Qualitative data from observations showed that teachers facilitated rich discussions about shape features, including corrective feedback, prompting comparisons, and group collaboration.

In the FG-EGP, teachers did not use formal mathematical measures such as degrees when defining or comparing shapes, as these concepts would be developmentally inappropriate for 3- to 4-year-old children. Instead, they emphasized qualitative and perceptual features of shapes through concrete, hands-on exploration. Teachers described and compared shapes using accessible language such as *edges*, *lines*, and *corners* (often substituting *vertices* with "corners").

During instruction, children were encouraged to trace the outlines of shapes with their fingers, imitate shapes using their hands or bodies, and draw them to reinforce recognition of their properties. Teachers prompted comparison by asking questions like, *"What is different between these two shapes?"* or *"Which one looks wider or thinner?"* For example, when distinguishing between a thin isosceles triangle and an equilateral triangle, teachers guided children to notice and describe differences in the openness or narrowness of lines rather than referring to angular measures. Children often expressed these observations in embodied terms, saying things like *"this one is fat"* or demonstrating with their hands that one triangle's base was wider than the other. Through this approach, the concept of shape was constructed experientially and visually rather than through formal geometry language. The focus was on developing spatial reasoning and comparative thinking, recognizing properties such as side length, orientation, and relative width, laying the foundation for later understanding of more abstract geometric measures such as angles and degrees.

This instructional discourse aligns with research by Christie and Gentner (2010) and Kotovsky and Gentner (1996), supporting the idea that conceptual alignment and comparison foster relational abstraction. The intervention addressed this barrier through structured teacher–child dialogue and Froebelian guided discovery. Teachers used intentional language and modeling strategies to help children differentiate between typical, atypical, and non-examples of shapes.

For example, in Klara's class, shape-based discussions were often personalized to emphasize identity and uniqueness: "*We talked about how each of them have a particular name and that they wouldn't want us to just call them student or child. They would like us to call them by their name. So that the shapes would like them to be called by their name. They just don't want to be called shape because every shape is different, just like every child is different.*"

Similarly, Charlotte emphasized precise mathematical vocabulary and corrected misconceptions through dialogue:

“*Yes, like I say, like I do tell them, yes, the shape has three sides. Let's count them. And we go on to other ones. They’ll say ones that are different.*”

She added: “*Ohh yes, I did say. This is the proper term for a diamond. It's a diamond shape, but the name is, the proper name is rhombus.*”

These examples reflect intentional teacher-led opportunities for concept comparison, guided language correction, and structural alignment of shape categories. In line with Christie and Gentner’s (2010) emphasis on relational abstraction through aligned comparisons, the FG-EGP provided systematic discourse that allowed children to notice and articulate defining properties, rather than rely solely on perceptual features. Beyond discussion, children’s hands-on engagement with Froebel Gifts played a pivotal role in consolidating geometric understanding. The following section explores how manipulation and physical interaction supported abstraction through sensory experience.

Addressing Barrier 3: Strengthening Children’s Focus on Defining Shape Features Through Hands-On Engagement

Although statistical gains in property-based language were modest, gesture-based reasoning significantly increased in the experimental groups. These embodied responses (e.g., counting corners, tracing edges) show that children engaged sensory-motorically with shape attributes, supporting abstraction even when verbal articulation was limited.

This suggests a shift toward integrated-concrete knowledge (Sarama & Clements, 2016), where children begin connecting visual input with physical manipulation and emerging symbolic understanding. The intervention leveraged Froebel Gifts to provide hands-on de-composition and composition tasks that supported this progression. Evidence from the teacher interviews further substantiates this. Charlotte (8-week experimental group teacher) stated that

> *One activity is, I gave them a bunch of, I set up sticks, and I wanted them to make a house out of the sticks and they did, and then they counted the sides and then they counted the corners and I also told them that, I said it but they knew too, that the corners, like they said it holds it together*.

She also added:

> *They also pointed to the top and the bottom. They always also used like, pointing like, this is a corner, that's the line. They would also point or, I would say "can you show me a line? Can you show me a corner?" in the shape and they would identify it by pointing it to me*.

Gisele (8-week experimental group teacher) noted "*Yes, with some shapes like a triangle and a square. I saw how they counted them*." These excerpts show children using their hands and bodies to locate and articulate shape features, often bridging informal vocabulary ("corners," "lines") with more formal instruction. This embodiment of geometric reasoning aligns with the aims of Froebel-based intervention and supports integrated-concrete understanding.

The FG-EGP provided rich opportunities for children to develop abstraction skills by facilitating their recognition of similarities and differences among geometric shapes, an essential first step in concept formation (Dienes, 1971; as cited in Gningue, 2006; Martinez & Huang, 2011). Teachers guided these processes through structured classroom dialogue. For instance, children were prompted to compare sides and lengths: "*T: Count the sides of the rectangle. How is it different from the triangle?*" and a child responded, "*C: The rectangle has more!*" (ON, 8-week experimental group - Destiny - April 25, 2024). These interactions encouraged attention to defining properties, steering children away from mere visual matching. Similarly, prompts like "*T: Oh, that one has shorter sides. You need one with long sides*" and "*You may look for one that has equal sides*" reflect a deliberate shift toward property-based reasoning (ON, 8-week experimental group - Charlotte - May 15, 2024).

Teachers also supported the abstraction process by engaging children in classifying and constructing composite shapes. Destiny, a teacher in the 4-week experimental group, observed how students began manipulating and categorizing shapes conceptually: "*They started to realize that they could build up with them too, which was kind of cool. But just being able to fix them and shape them into maybe taking two squares and making a rectangle now.*" She further noted that "*Especially when we talked about the different triangles... they were actually like quite smart about it. Like, oh, these two sides are longer.*" Such activities align with Dienes's cycle of abstraction, in which structured play and comparison support concept generalization.

In addition, teachers explicitly incorporated shape properties into their instructional language. Grace explained, "*We talked about how many sides and we talked about, you know, whether it was smooth,*" and added, "*Sometimes when we talk about the square, we talk about corners. How many corners does it have?*" Destiny similarly described: "*We had a period where we were asking them to find different shapes and you know, talking about their surfaces and whatnot.*" These teacher–child dialogues provided repeated opportunities to abstract defining features from shape categories, aligning with the constructivist view that meaning is built through reflection on properties (Mitchelmore & White, 2007).

Children also participated in activities involving composition and de-composition, which supported mental manipulation of geometric forms. For example, during a Froebel Gift 7 activity, the teacher prompted, "*Try to make a square using these triangles... Now it's a square!*"

and a child confirmed, "*Yes!*" The teacher then reinforced the concept: "*The two triangles make one big square!*" (ON, 8-week experimental group - Klara - May 17, 2024). Similarly, shape transformation tasks further engaged students in reflective abstraction: "*T: What happens if we rotate this rectangle?*" "*C: It still looks the same!*" "*T: What if we rotate the triangle?*" "*C: Now the point is here!*" (ON, 8-week experimental group - Gisele - May 17, 2024). These classroom interactions exemplify how hands-on manipulation and verbal reflection work together to promote abstraction, a process that extends beyond physical activity to the mental reconstruction of experiences, as discussed in broader theoretical perspectives on reflective abstraction and conceptual development. The children's dialogue and actions in these excerpts illustrate the very processes described by scholars such as Cetin and Dubinsky (2017) and Clements (2000), where abstraction emerges through guided reflection on concrete experiences rather than through manipulation alone.

Cetin and Dubinsky (2017) apply the concept of reflective abstraction to computational thinking, illustrating how learners abstract and generalize from specific experiences. It underscores the importance of reflection in developing higher-order thinking skills. Clements (2000) critiques the assumption that manipulatives inherently convey mathematical meaning. He argues that without intentional instruction and reflection, manipulatives may not effectively support conceptual understanding. Taken together, these perspectives highlight that meaningful learning with manipulatives depends on purposeful guidance and reflection, a principle that also underpins Froebel's philosophy and is clearly echoed in the design and outcomes of the FG-EGP.

My findings resonate strongly with Froebel's original pedagogical philosophy and its modern interpretations. Reinhold et al. (2017) emphasize that the Froebel Gifts were designed to help children discern mathematical structure through play, composition, and variation, principles reflected in the FG-EGP's use of comparison, de-composition, and reconstruction tasks. Similarly, Tovey (2019) highlights that Froebelian practice remains highly relevant today because it links hands-on exploration with guided dialogue, allowing children to move from perceptual engagement toward conceptual understanding. This connection is evident in the 8-week experimental group, where sustained manipulation and teacher-scaffolded discussion of shapes led to increased property-based reasoning and recognition of atypical forms.

The integration of these materials also echoes Correia and Fisher's (2014) argument that Froebel Gifts strengthen conceptual understanding by making abstract relationships visible through concrete manipulation. As Brosterman (1997) observes, Froebel's sequence of Gifts was intentionally geometric, progressing from spheres to cubes to tablets, to cultivate awareness of form, symmetry, and transformation. Within the FG-EGP, this same progression enabled children to experience geometric properties through play and language simultaneously, bridging sensory activity and reflective abstraction. Collectively, this literature supports the interpretation that the FG-EGP's effectiveness stems not from the materials alone, but from their Froebelian integration of structured variation, teacher guidance, and dialogic exploration that transforms tactile play into conceptual insight.

Gesture-based responses also played a critical role in this developmental shift. Teachers encouraged children to physically manipulate shapes: "*T: Turn your shapes; maybe it will be different*" (ON, 8-week experimental group - Gisele - May 17, 2024) and students demonstrated embodied understanding with constructions like, "*C: I made a circle tower*" (ON, 4-week experimental group - Destiny - April 22, 2024). These sensorimotor engagements provided an essential bridge between visual perception and symbolic reasoning, marking early stages of integrated-concrete knowledge (Sarama & Clements, 2016). As such, the FG-EGP effectively supported the early stages of abstraction by combining verbal scaffolding, manipulative-based composition and de-composition, and gesture-driven reflection, all core processes identified by Piaget and Dienes, and in constructivist theories of abstraction. Taken together, these pedagogical insights set the stage for a theoretical interpretation of the findings. The next section connects the observed learning patterns to major cognitive and constructivist frameworks that explain how abstraction develops in early geometry learning.

Implications

The findings of this study carry significant implications for theory, curriculum, and future research in early geometry education. Collectively, they demonstrate that abstraction in young children's geometric reasoning emerges through the interplay of cognitive, embodied, and social processes. The study extends existing theoretical frameworks, particularly those of Piaget, Dienes, Fischbein, and Vinner, by revealing how guided manipulation, gesture, and comparison can bridge perceptual understanding and conceptual generalization. At the curricular level, the FG-EGP illustrates how structured play and teacher-guided dialogue can be integrated within Ontario's play-based frameworks to promote property-based reasoning without sacrificing developmental appropriateness. Finally, the findings open new directions for research, emphasizing the need to examine gesture, embodiment, and duration of intervention as critical variables in fostering abstraction. Together, these implications affirm that young children's geometric learning is not solely a cognitive process but a recursive, embodied, and dialogic construction of meaning that connects theory, pedagogy, and future inquiry. The following subsections elaborate on these implications in relation to theoretical perspectives, curriculum design, and future research directions. Building on the theoretical frameworks discussed earlier, the first implication centres on how this study contributes to and extends existing theories of abstraction by integrating cognitive, constructivist, and embodied perspectives.

Implications for Theory

The results show that children in the 8-week experimental group not only performed better on shape recognition tasks but also demonstrated increased use of gestures, manipulative engagement, and some property-based reasoning. These findings suggest a transition from empirical abstraction which is based on perceptual cues and physical interaction (e.g., gestures, manipulatives) toward the beginnings of reflective abstraction, where children mentally manipulate and reorganize knowledge. This trajectory aligns with Piaget's theory, in which abstraction is constructed actively rather than passively absorbed.

The findings of this study reflect the dynamic and multifaceted nature of abstraction, as conceptualized by Piaget, Dienes, and Mitchelmore and White, and more recently, by theorists of embodied cognition such as Maturana and Varela (1987). Piaget's distinction between empirical and reflective abstraction helps explain the shift observed in the experimental groups. While children began by recognizing shapes through sensory features (empirical abstraction), the intervention promoted a gradual internalization of properties through comparison and manipulation, laying the groundwork for reflective abstraction. Dienes's theory, particularly his six-stage model and three-cycle learning progression, offers further insight into how structured play, comparison tasks, and guided manipulation of Froebel Gifts scaffold children's abstraction processes. According to Dienes, abstraction arises through recognizing commonalities across varied perceptual contexts, a process supported by FG-EGP 's use of typical, atypical, and non-examples. Children were exposed to multiple embodiments of shapes (e.g., rods, beads, Play-Doh, planes, 3D objects), which promoted mental generalization through sensory variation which is critical for flexible abstraction (Dienes, 1971; as cited in Gningue, 2006). My findings that children began to notice that “two triangles make a square” or that “spheres roll but cubes don’t” reflect Dienes’s notion that abstraction emerges when children notice invariants across varying representations, thereby forming robust mental prototypes grounded in conceptual understanding. Building on these perspectives of abstraction as a process of recognizing invariants across experiences, Vinner’s distinction between concept image and concept definition offers a complementary lens for understanding how children’s emerging geometric reasoning transitions from intuitive, perceptual recognition to more analytical, property-based understanding.

Vinner’s distinction between concept image (prototype-based, intuitive) and concept definition (analytic, formal) is crucial for interpreting my results. My findings show a shift from System 1 (prototype-based thinking) toward System 2 (analytical, property-based reasoning) among children in the 8-week experimental group, albeit modest and not always statistically significant. The frequent gesture use implies children were bridging visual images and emerging understanding of properties (e.g., counting corners), supporting the idea that concept images are in transition. Cognitive load theory helps explain why many children didn’t reach statistically significant gains in verbal property-based responses: working memory demands for abstract geometric definitions may still exceed young children's capacity, especially when it is not scaffolded. While Vinner’s framework illuminates the cognitive transition from intuitive to analytical reasoning, Fischbein’s theory extends this understanding by emphasizing that geometric thinking simultaneously engages both visual imagery and conceptual reasoning, a duality that is strongly reflected in my participants’ evolving shape understandings.

Fischbein's theory that a geometric concept is both visual and conceptual is perhaps most directly affirmed by my findings. Children's manipulation and verbalization of shape properties show that both sensory imagery and formal properties were in use, the essence of a figural concept. The shift from "*this looks like a box*" to "*it has four sides and four corners*" indicates a movement from static figural concepts to dynamic ones, consistent with the framework proposed by Walcott et al. (2000). The improvement in recognizing skewed or rotated shapes also supports conceptual change theory: children are gradually revising prior misconceptions (e.g., "triangles must point up") and refining their figural concepts. This gradual evolution underscores neuro-constructivist views that conceptual growth is driven by experience and interaction, leading to reorganization of mental and neural representations over time (Mareschal et al., 2007).

Moreover, Mitchelmore and White's (2000; 2007) "Teaching for Abstraction" model aligns closely with this progression in a way that children first encountered shape varieties, then compared them, abstracted key properties, and began generalizing this understanding to new shapes. Mitchelmore and White's (2000; 2007) four-stage model, experiencing, comparing, reifying, and applying mirrors my instructional strategy and data trajectory. Children experienced varied shapes and manipulations through Froebel Gifts. They compared properties through structured comparison tasks. Some began to reify shared properties like "sides" and "vertices." A few even applied these abstractions, using them in shape construction and verbal explanation. The qualitative findings (e.g., identifying squares through counting corners or decomposing shapes into components) provide evidence that children were not just recognizing shapes but beginning to generalize properties across contexts, a key tenet of both abstraction and situated learning. The embodied actions observed reinforce the idea that learning is grounded in action. The children's reliance on gesture and manipulation prior to verbal explanation, from one lens, may suggest that bodily enactments are precursors to abstract thought (Goldin-Meadow, 2014; Núñez, 2000). As Goldin-Meadow (2014) argues, gestures serve as a bridge between action and thought, revealing ideas that are not yet fully verbalized but are actively taking shape in the learner's mind. These embodied movements externalize cognition, allowing children to "think with their hands" before formalizing concepts in language. Similarly, Núñez (2000) contends that mathematical understanding is rooted in bodily experience, as abstract concepts emerge through sensorimotor schemas and embodied metaphors that structure reasoning. From this perspective, the children's gestures and manipulations during the activities can be viewed not as peripheral behaviors but as essential cognitive acts that mediate the transition from concrete experience to abstract geometric understanding.

However, these cognitive theories alone do not fully account for the embodied dimension of learning evidenced in my study, particularly the statistically significant increase in gesture-based responses, such as children tracing edges or pointing to vertices. These findings resonate with Maturana and Varela's enactive view of cognition, which argues that knowledge emerges through recursive interactions between the organism and its environment (Maturana & Varela, 1987).

When I began this study, I did not anticipate that children would frequently use gestures as part of their reasoning during the post-test. Consequently, I did not initially consider Maturana and Varela's framework of embodied cognition as a guiding theoretical lens, and their work was not elaborated upon in my literature review. My primary theoretical framework was grounded in Fischbein's figural concept, which was well aligned with the study's original focus on the interplay between visual and conceptual understanding in geometric reasoning. However, during the analysis, gesture-based responses emerged as a recurrent and meaningful pattern, suggesting that children's abstraction processes were not only cognitive but also deeply embodied. This unanticipated finding highlights an important direction for future research. In subsequent studies, I plan to incorporate Maturana and Varela's (1987) perspective more substantively to explore how sensorimotor experiences, gestures, and embodied interactions contribute to the development of geometric abstraction in early childhood.

From an enactivist perspective, abstraction is not treated as a mental representation detached from action, but as a form of sensorimotor coordination through which a learner *"brings forth"* a world of meaning, that is, actively constitutes or enacts meaning through interaction rather than passively receiving it (Maturana & Varela, 1987). Children's bodily engagement with Froebel Gifts, through composing, decomposing, comparing, and manipulating shape components, illustrates how abstract thought emerges through *structural coupling*, defined as the reciprocal, ongoing interaction and mutual shaping between the learner and the learning environment (Maturana & Varela, 1987). Their manipulation of geometric components instantiated the kind of recursive coordination that Maturana and Varela (1987) describe as foundational to cognition. From this standpoint, gesture-based abstraction is not merely a developmental precursor to verbal reasoning but a *co-equal pathway*—an equally valid and knowledge-generating mode of mathematical sense-making alongside linguistic expression (see also Alibali & Nathan, 2012).

Together, these perspectives suggest that abstraction in early geometry learning is not a linear transition from sensory to symbolic, but a recursive, embodied process in which sensorimotor actions, mental reflection, and language co-evolve (Clements & Sarama, 2000; Clements et al., 1999) also highlights that children's understanding of shapes progresses from recognizing prototypical examples to identifying shapes based on properties. They emphasize the importance of guided experiences and discussions to facilitate this development. My findings affirm that extended, hands-on geometry instruction anchored in the Froebel approach and Gifts and enriched with comparison, composition, and guided discussion can foster this developmental trajectory.

The gesture-based responses, although categorized as “visual,” actually support neuro-constructivist principles: learning as a progressive, embodied reorganization of neural structures. Gestures serve as embodied scaffolds bridging sensory experiences and mental operations (Alibali & Nathan, 2012), suggesting early internalization of shape properties even when children are still developing the language to express them. This developmental trajectory also affirms the constructivist contention that knowledge is actively built through interaction with the environment. Children’s increasing tendency to manipulate and explore shape properties indicates they are constructing new mental representations through Froebel Gifts and tasks that shift attention from superficial to structural features.

My findings collectively affirm that abstraction in early childhood geometry can emerge incrementally from embodied, intuitive action to partial verbalization and conceptual flexibility. The Froebel-based instruction appeared particularly effective in supporting this shift by:

- Offering multiple representations
- Structuring guided comparisons
- Engaging children in embodied learning
- Facilitating movement from prototype-based recognition to property-based reasoning

This integrated analysis not only aligns with my theoretical framework but also expands upon it by connecting each result strand to broader cognitive, developmental, and instructional paradigms. It validates the FG-EGP’s potential to lay a cognitive foundation for abstraction through structured, hands-on, and dialogic engagement. These theoretical connections also carry significant implications for early childhood curriculum design. The following section situates the FG-EGP within Ontario's curriculum frameworks, showing how its principles align with and enhance play-based and inquiry-driven approaches. Moving from theory to practice, the next subsection examines how the principles of the FG-EGP can inform curriculum design and early childhood pedagogy, aligning theoretical insights with practical applications in play-based learning environments.

Implications for Curriculum Design

The findings of this study align meaningfully with the pedagogical goals outlined in *Early Learning for Every Child Today* (ELECT; Best Start Expert Panel on Early Learning, 2007), the foundational curriculum framework for early childhood education in Ontario. ELECT emphasizes play-based, inquiry-driven learning in naturalistic classroom environments, where children's cognitive, social, and physical development are nurtured through rich, open-ended experiences. The FG-EGP, implemented within this context, complements and extends the goals of ELECT by structuring children's exploratory play into guided geometric experiences that foster both perceptual and conceptual learning. Rather than replacing emergent, child-centred learning, the FG-EGP supports it through intentional, scaffolded engagements that draw attention to mathematical relationships and shape properties.

One of the key strengths of the FG-EGP is that it fits within the existing rhythms of early learning classrooms, where daily routines already include classification, construction, measurement, and visual-spatial reasoning. Activities such as sorting blocks, building towers, and comparing sizes are common across the daycares participating in this study. However, the FG-EGP adds a theoretical backbone and pedagogical intent to these activities, transforming them into opportunities for abstraction. By focusing on typical and atypical examples of shapes and embedding geometry vocabulary into hands-on tasks, the program helped children move from intuitive recognition to a more analytical, property-based understanding of shape categories, an outcome fully compatible with the goals of ELECT's emphasis on cognitive development through play.

Furthermore, the FG-EGP's emphasis on teacher–child dialogue and gesture-based exploration aligns with Vygotsky's sociocultural theory and the ELECT framework's focus on co-constructed learning within the child's zone of proximal development. ELECT encourages educators to act as co-learners and co-researchers, engaging with children's inquiries and extending their thinking through open-ended questioning and discussion. This role was made explicit in the FG-EGP, where teachers were trained to use guided comparison, precise vocabulary, and Froebel Gifts to support children's developing geometric concepts. The program encouraged children not only to name shapes but to articulate reasons for their classifications, an instructional move that fostered both language development and mathematical reasoning.

Additionally, the FG-EGP builds on the broader educational aims outlined in Ontario's *Kindergarten Program* (Government of Ontario, 2016), which stresses the importance of problem-solving, visual-spatial exploration, and the use of concrete materials to build foundational mathematical concepts. Within this framework, mathematics is not taught as a discrete subject but is interwoven into children's explorations of their environment. The structured yet playful nature of FG-EGP fits seamlessly into this pedagogical approach, helping children connect their everyday experiences with more abstract mathematical ideas. The integration of Froebel Gifts and spatial reasoning tasks in FG-EGP supports the development of operational knowledge, enabling children to physically manipulate and internalize geometric relationships—a goal shared by both ELECT and *The Kindergarten Program*.

Importantly, this study also highlights a potential gap in current curriculum practices: while ELECT and similar frameworks encourage play-based mathematics, they may not always offer clear strategies for fostering abstraction, particularly in shape recognition. The results suggest that without structured comparison, guided discourse, and intentional exposure to non-prototypical shapes, children may default to surface-level features in geometry tasks. The FG-EGP serves as a model for how educators can retain the developmental appropriateness of play while incorporating explicit pedagogical supports that guide children toward recognizing invariant properties across shape examples. This has significant implications for professional development and curriculum design in early childhood education, suggesting that teachers need more support (e.g., professional learning in structured comparison strategies, mathematical discourse facilitation, and manipulative-based geometry instruction) in bridging play-based learning with cognitively rich instruction.

In sum, the FG-EGP complements and strengthens existing curriculum frameworks by offering a developmentally appropriate yet theoretically grounded approach to geometry instruction. It capitalizes on the strengths of play-based learning environments while addressing the limitations of unstructured exploration in fostering geometric abstraction. By integrating manipulative use, gesture, teacher–child dialogue, and structured comparison, the FG-EGP offers a practical pathway for early childhood educators to deepen children's spatial reasoning and conceptual understanding of shape. These connections affirm the program's relevance and scalability within the broader landscape of early childhood education in Ontario and beyond. Finally, building on these curricular implications, the last section outlines directions for future research, identifying how subsequent studies can further advance understanding of abstraction, embodiment, and instructional design in early geometry education. Building on the theoretical and curricular insights discussed above, the final subsection outlines directions for future research, highlighting how further investigation into embodiment, gesture, and instructional design can deepen understanding of abstraction and refine early geometry education practices.

Implications for Future Research

The findings of this study open multiple avenues for future research, particularly around how young children develop abstract geometric reasoning through embodied, manipulative-rich instruction. Given the observed shift from empirical to emerging reflective abstraction, subsequent studies could examine how different stages of abstraction unfold over time and how various instructional modalities (e.g., gesture-rich environments, spatial language scaffolding, digital manipulative tools) accelerate or mediate this developmental trajectory. Longitudinal designs following children beyond the preschool years would be especially valuable in determining whether early gesture-based or manipulative interactions predict later success in formal geometric reasoning or other spatial-mathematical domains.

Future research might also investigate more deeply the role of gesture as both a diagnostic and pedagogical tool. While gesture was categorized in this study under "visual" responses, its significance as a developmental bridge toward verbal and conceptual abstraction merits independent analysis. Building on the work of Alibali and Nathan (2012) and Goldin-Meadow (2014), future studies could code and differentiate between types of gestures (e.g., deictic, iconic, metaphoric) and analyze their correspondence with stages of abstraction. Experimental designs could test whether encouraging gesture production during geometry tasks enhances children’s ability to verbalize and generalize geometric properties, or whether certain gestures serve as predictors of conceptual breakthroughs.

The embodied dimension of learning also presents rich ground for further exploration. Research drawing on enactivist and neuro-constructivist frameworks (Mareschal et al., 2007; Maturana & Varela, 1987) could investigate how sensorimotor experiences with physical manipulatives facilitate structural reorganization of spatial and mathematical thinking. Neurocognitive methods, such as eye tracking or EEG, might provide empirical support for the embodied pathways proposed in this study, revealing how neural activity shifts as children move from manipulation to mental imagery and verbal explanation. Investigating how bodily engagement differentially supports learners with varying developmental profiles or neurodiversity could also enrich our understanding of inclusive geometry instruction.

Another critical direction for future research lies in refining program design. The differences observed between the 4-week and 8-week implementations suggest that intervention duration and instructional density play significant roles in fostering abstraction. Future studies should investigate optimal timing and sequencing of geometry activities across varied age groups and educational contexts. Comparing the FG-EGP model with other instructional frameworks (e.g., Reggio Emilia, Montessori, spatial reasoning programs) could help identify the most effective strategies for engaging young learners in abstract geometric thinking. Research should also consider how teachers' professional development influences the fidelity and quality of abstraction-centred geometry instruction.

Additionally, integration of language and mathematics warrants further exploration. While some children in this study demonstrated conceptual understanding through gesture, others were still developing the language needed to verbalize their insights. Investigating how targeted language supports such as mathematics discourse routines, vocabulary scaffolds, or bilingual instruction, interact with embodied learning could yield valuable insights into how verbal and nonverbal pathways support one another. The role of teacher questioning and discourse strategies in facilitating transitions from gesture to speech-based abstraction is a particularly fertile area for inquiry.

Finally, mixed methods approach, like the one employed in this study, should continue to be refined and expanded in future research. By combining statistical measures with rich observational and interview data, researchers can better capture the nuanced, gradual nature of abstraction development in early childhood. Future studies may also benefit from developing more sensitive coding tools to distinguish between emerging conceptual reasoning, rote language use, and embodied understanding. The findings of this study point to a layered developmental process in which abstraction arises not solely through language or logic, but through recursive, embodied interactions (Maturana & Varela, 1987) with objects, people, and environments, an insight that must guide both future research and instructional design.

In this chapter, I discussed how the Froebel Gifts Early Geometry Program (FG-EGP) influenced 3- to 4-year-old children’s abstraction processes in geometric shape recognition by integrating quantitative and qualitative evidence to offer a holistic interpretation of the findings. Across the research questions, a consistent pattern emerged: when children experienced sustained, structured opportunities to compare, compose, and decompose shapes, supported by intentional teacher discourse and Froebel Gifts, they were more likely to move beyond prototype-based recognition toward attention to invariant properties. Quantitative results showed the strongest gains in the 8-week experimental group, underscoring the importance of duration for concept formation, while qualitative data revealed how teachers’ vocabulary, questioning, corrective feedback, and comparison routines shaped children’s reasoning in real time. Although increases in property-based verbal responses were mixed across the full sample, the qualitative findings and the significant rise in gesture-based reasoning suggest that abstraction often developed through embodied pathways before it could be consistently expressed in formal language. Together, these results extend and connect Piaget's and Dienes's accounts of abstraction with Fischbein and Vinner's work on figural concepts and concept images, while also highlighting embodiment as a meaningful route for early geometric thinking. Overall, the study demonstrates that developmentally appropriate play can remain central in early childhood classrooms while becoming cognitively rich when paired with structured variation, guided discourse, and purposeful manipulative use. These conclusions carry direct implications for curriculum design and professional learning: to support young children's abstraction, teachers benefit from practical tools and training that help them bridge play-based exploration with intentional, property-focused instruction.

References

Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences, 25*(2), 203–239. https://doi.org/10.1080/10508406.2016.1143370

Ahmetoğlu, E., & İlhan İldiz, G. (2018). The Friedrich Froebel approach. In R. Efe, I. Koleva, & E. Atasoy (Eds.), *Recent researches in education* (pp. 355–366). Cambridge Scholars Publishing. https://www.researchgate.net/publication/328018784_The_Friedrich_Froebel_Approach

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*(2), 247-286. https://doi.org/10.1080/10508406.2011.611446

Arnas, Y. A., & Aslan, D. (2005). Geometry in early childhood. *Eğitim Bilim Toplum Dergisi*, *3*, 36–45.

Arnas, Y. A., & Aslan, D. (2007). Content analysis of the presentation of geometric shapes in preschool education materials. *Journal of Çukurova University Institute of Social Sciences*, *16*, 69–80.

Arnas, Y. A., & Aslan, D. (2010). Children's classification of geometric shapes. *Journal of Çukurova University Institute of Social Sciences*, *1*(19), 254–270.

Arshavsky, N., & Goldenberg, E. P. (2005). Perceptions of a quadrilateral in a dynamic environment. In D. Carraher, R. Nemirovsky, & A. D. Schliemann (Eds.), *Medium and meaning: Video papers in mathematics education research* (*Journal for Research in Mathematics Education* Monograph 13). National Council of Teachers of Mathematics.

Aslan, D., & Arnas, Y. A. (2007a). A study into the effects of teaching geometric shapes to preschool children through drama. *Hacettepe University Journal of Education*, *32*, 35–44.

Aslan, D., & Arnas, Y. A. (2007b). Three- to six-year-old children's recognition of geometric shapes. *International Journal of Early Years Education, 15*(1), 83–104. https://doi.org/10.1080/09669760601106646

Bartolini Bussi, M. G., & Mariotti, M. A. (2008). Semiotic mediation in the mathematics classroom: Artifacts and signs after a Vygotskian perspective. In L. D. English, M. G. Bartolini Bussi, G. H. Jones, R. A. Lesh, B. Sriraman, & D. Tirosh (Eds.), *Handbook of International Research in Mathematics Education* (2nd ed., pp. 746-783). Routledge.

Benedek, A. G. (2018). *Embodied conceptions of mathematical understanding in the twentieth century: The emergence of Zoltán P. Dienes's principles and their origin.* Paper presented at the Tenth HMTM: History of Mathematics and Teaching of Mathematics Conference, University of Miskolc, Miskolc, Hungary. https://real.mtak.hu/80683/1/Benedek_Paper_A4_format_v13_u.pdf

Best Start Expert Panel on Early Learning. (2007, January). *Early Learning for Every Child Today: A framework for Ontario early childhood settings*. https://tinyurl.com/237zuhp2

Borasi, R. (1984). Some reflections on and criticisms of the principle of learning concepts by abstraction. *For the Learning of Mathematics, 4*(3), 14–18. http://flm-journal.org/Articles/10B18404086822C4FD9B1768D858E0.pdf

Boswell, D. A., & Green, J. F. (1982). The abstraction and recognition of prototypes by children and adults. *Child Development, 53*(4), 1028-1037. https://doi.org/10.2307/1129144

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

Brosterman, N. (1997). *Inventing kindergarten.* Harry N. Abrams.

Burger, W. F., & Shaughnessy, J. M. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for Research in Mathematics Education, 17*(1), 31–48. https://doi.org/10.2307/749317

Burns, A. (1999). *Collaborative action research for English language teachers*. Cambridge University Press.

Cetin, I., & Dubinsky, E. (2017). Reflective abstraction in computational thinking. *Journal of Mathematical Behavior*, *47*, 70–80. https://doi.org/10.1016/j.jmathb.2017.06.004

Charlesworth, R. (2005). Prekindergarten mathematics: Connecting with national standards. *Early Childhood Education Journal*, *32*(4), 229–236. https://doi.org/10.1007/s10643-004-1423-7

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, *11*(3), 356–373. https://doi.org/10.1080/15248371003700015

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*(1), 43-71. https://doi.org/10.1016/0010-0277(86)90053-3

Clements, D. H. (2000). "Concrete" manipulatives, concrete ideas. *Contemporary Issues in Early Childhood*, *1*(1), 45–60. https://doi.org/10.2304/ciec.2000.1.1.7

Clements, D. H. (2001). Mathematics in the preschool. *Teaching Children Mathematics*, *7*(5), 270–275. https://doi.org/10.5951/TCM.7.5.0270

Clements, D. H. (2003). *Teaching and learning geometry*. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 151-178). National Council of Teachers of Mathematics.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420-464). Macmillan.

Clements, D. H., DiBiase, A.-M., & Sarama, J. (Eds.). (2004). *Engaging young children in mathematics: Standards for early childhood mathematics education*. Lawrence Erlbaum Associates.

Clements, D. H., & Sarama, J. (2000). Young children's ideas about geometric shapes. *Teaching Children Mathematics*, *6*(8), 482–488. https://doi.org/10.5951/TCM.6.8.0482

Clements, D. H., & Sarama, J. (2011). Early childhood teacher education: The case of geometry. *Journal of Mathematics Teacher Education*, *14*(2), 133–148. https://doi.org/10.1007/s10857-011-9173-0

Clements, D. H., & Sarama, J. (2020). *Learning and teaching early math: The learning trajectories approach* (3rd ed.). Routledge.

Clements, D. H., Sarama, J., & Joswick, C. (2022). Learning and teaching geometry in early childhood. In A. Sharif-Rasslan & D. Hassidov (Eds.), *Special issues in early childhood mathematics education research: Learning, teaching and thinking* (pp. 95–131). Brill. https://doi.org/10.1163/9789004510685_005

Clements, D. H., Sarama, J., & Liu, X. H. (2008). Development of a measure of early mathematics achievement using the Rasch model: The Research-Based Early Maths

Assessment. *Educational Psychology*, *28*(4), 457–482. https://doi.org/10.1080/01443410701777272

Clements, D. H., Sarama, J., Swaminathan, S., Weber, D., & Trawick-Smith, J. (2018). Learning and teaching geometry in early childhood. *Quadrante, 27*(2), 7–32. https://www.researchgate.net/publication/330092761_Learning_and_teaching_geometry_in_early_childhood

Clements, D. H., Swaminathan, S., Hannibal, M. A. Z., & Sarama, J. (1999). Young children's concepts of shape. *Journal for Research in Mathematics Education*, *30*(2), 192–212. https://doi.org/10.2307/749610

Copley, J. V. (2000). Geometry and spatial sense in the early childhood curriculum. In *The young child and mathematics* (pp. 105–124). National Association for the Education of Young Children.

Correia, J., & Fisher, M. H. (2014). Froebel Gifts: A tool to reinforce conceptual knowledge of fractions. *Ohio Journal of School Mathematics, 69*, 31–36. https://tinyurl.com/bdhs67tc

Creswell, J. W. (2006). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage Publications.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.

Cross, C. T., Woods, T. A., & Schweingruber, H. (eds.). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. National Academies Press. https://doi.org/10.17226/12519

Dauksas, L., & White, J. (2014, April/May). Discovering shapes and space in preschool. *Teaching Young Children, 7*(4). National Association for the Education of Young

Children. https://www.naeyc.org/resources/pubs/tyc/apr2014/discovering-shapes-and-space-preschool

Dienes, Z. P. (1971). *Building up mathematics* (4th ed.). Hutchinson Educational Ltd.

Downton, A., Livy, S., & Reinhold, S. (2019). Grade 3/4 students' understanding of geometrical objects: Australian case studies on (mis)conceptions of cubes. In U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 857–864). ERME.

Dubinsky, E. (2002). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 95–126). Springer. https://doi.org/10.1007/0-306-47203-1_7

Duncan, R. M., & Tarulli, D. (2003). Play as the leading activity of the preschool period: Insights from Vygotsky, Leont'ev, and Bakhtin. *Early Education and Development, 14*(3), 271–292. https://doi.org/10.1207/s15566935eed1403_2

Eugene, F., Jr., & Provenzo, J. M., Jr. (2009). Friedrich Froebel's gifts: Connecting the spiritual and aesthetic to the real world of play and learning. *American Journal of Play, 2*(1), 85–99.

Evans, C. (2018). *Analysing semi-structured interviews using thematic analysis: Exploring voluntary civic participation among adults* [Dataset]. SAGE Publications Ltd. https://doi.org/10.4135/9781526439284

Ferrara, K., Hirsh-Pasek, K., Newcombe, N., Golinkoff, R., & Lam, W. (2011). *Block talk: Spatial language during block play. Mind, Brain, and Education*, 5(3), 151-163. https://doi.org/10.1111/j.1751-228X.2011.01122.x

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. *Health Services Research*, *48*(6), 2134–2156. https://doi.org/10.1111/1475-6773.12117

Fischbein, E. (1993). The theory of figural concepts. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 139–147). Lawrence Erlbaum Associates.

Fischbein, E., & Nachlieli, T. (1998). Concepts and figures in geometrical reasoning. *International Journal of Science Education, 20*(10), 1193–1211. https://doi.org/10.1080/0950069980201003

Fisher, K. R., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Taking shape: Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development, 84*(6), 1872-1878. https://doi.org/10.1111/cdev.12091

Fitzpatrick, R. (Ed. & Trans.). (2008). *Euclid's elements of geometry*. https://farside.ph.utexas.edu/Books/Euclid/Elements.pdf

Fujita, T., & Jones, K. (2006). Primary trainee teachers' knowledge of parallelograms. *Proceedings of the British Society for Research into Learning Mathematics*, *26*(2), 25–30.

Gecu-Parmaksiz, Z., & Delialioglu, O. (2019). Augmented reality-based virtual manipulatives versus physical manipulatives for teaching geometric shapes to preschool children. *British Journal of Educational Technology, 50(*6), 3376–3390. https://doi.org/10.1111/bjet.12740

Gentner, D., & Goldin-Meadow, S. (Eds.). (2003). *Language in mind: Advances in the study of language and thought.* Boston Review. https://doi.org/10.7551/mitpress/4117.001.0001

Ginsburg, H. P., & Oppenzato, C. (2018). *What children know and need to learn about shape and space.* DREME/Stanford University. https://prek-math-te.stanford.edu/spatial-relations/what-children-know-and-need-learn-about-shape-and-space

Gningue, S. M. (2006). Students working within and between representations: An application of Dienes's variability principles. *For the Learning of Mathematics, 26*(2), 41-47. https://flm-journal.org/Articles/1D0AAA216C196F0F99E5287578BB0B.pdf

Gningue, S. M. (2016). *Remembering Zoltán P. Dienes, a maverick of mathematics teaching and learning: Applying the variability principles to teach algebra.* International Journal for Mathematics Teaching and Learning, 17 (2), 122–146. https://doi.org/10.4256/ijmtl.v17i2.17

Goldin-Meadow, S. (2014). How gesture works to change our minds. *Trends in Neuroscience and Education, 3*(1), 1-10. https://doi.org/10.1016/j.tine.2014.01.002

Gómez, R. L., & LaKusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7*(5), 567–580. https://doi.org/10.1111/j.1467-7687.2004.00381.x

Gordon, R., & Ramani, G. B. (2021). Embodied cognition and early mathematics learning: A framework and review of evidence. *Developmental Review*, *60*, Article 100960. https://doi.org/10.3389/fpsyg.2021.650286

Government of Ontario. (2016). *The kindergarten program*. https://www.ontario.ca/document/kindergarten-program-2016

Government of Ontario. (2020, June 23). *Ontario introduces new math curriculum for elementary school students* [Press release]. https://news.ontario.ca/en/backgrounder/57346/ontario-introduces-new-math-curriculum-for-elementary-school-students

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*(3), 280-290. https://doi.org/10.1016/j.jecp.2010.04.017

Gray, E., & Tall, D. (2001). Relationships between embodied objects and symbolic procepts: An explanatory theory of success and failure in mathematics. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of 25th Annual Conference of the International Group for the Psychology of Mathematics Education*, *3*, 65–72.

Gray, E., & Tall, D. (2007). Abstraction as a natural process of mental compression. *Mathematics Education Research Journal*, *19*(2), 23–40. https://doi.org/10.1007/BF03217454

Halat, E., & Yeşil Dağlı, Ü. (2016). Preschool students' understanding of a geometric shape, the square. *Bolema: Boletim de Educação Matemática*, *30*(55), 830–848. https://doi.org/10.1590/1980-4415v30n55a25

Hallowell, D. A., Okamoto, Y., Romo, L. F. & La Joy, J. R. (2015). First-graders' spatial-mathematical reasoning about plane and solid shapes and their representations. *The International Journal on Mathematics Education*, *47*(3), 363–375. https://doi.org/10.1007/s11858-015-0664-9

Hannibal, M. A. Z. (1999). Young children's developing understanding of geometric shapes. *Teaching Children Mathematics, 5*(6), 353–358.

Hannibal, M. A. Z., & Clements, D. H. (2010). *Young children's understanding of basic geometric shapes* [Unpublished manuscript]. National Science Foundation, Grant ESI-8954644. https://www.researchgate.net/publication/258933076_Young_children%27s_understanding_of_basic_geometric_shapes

Hawes, Z., & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychonomic Bulletin & Review*, *27*, 465–482. https://doi.org/10.3758/s13423-019-01694-7

Hershkowitz, R. (1989). Visualization in geometry—Two sides of the coin. *Focus on Learning Problems in Mathematics, 11*(1-2), 61–76

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). *A mandate for playful learning in preschool: Presenting the evidence.* Oxford University Press.

Jones, K. (1998). Theoretical frameworks for the learning of geometrical reasoning. *Proceedings of the British Society for Research Into Learning Mathematics*, *2*(18), 29–34. https://doi.org/10.1007/BF01273689

Jones, K. (2000). Teacher knowledge and professional development in geometry. *Proceedings of the British Society for Research into Learning Mathematics, 20*(3), 109-114.

Jones, K., Mooney, C., & Harries, T. (2002). Trainee primary teachers' knowledge of geometry for teaching. *Proceedings of the British Society for Research into Learning Mathematics*, *22(*2), 95–100. https://eprints.soton.ac.uk/12596/1/Jones_Mooney_Harries_BSRLM_22-2_2002.pdf

Kalénine, S., Pinet, L., & Gentaz, E. (2011). The visual and visuo-haptic exploration of geometrical shapes increases their recognition in preschoolers. *International Journal of Behavioral Development*, *35*, 18–26. https://doi.org/10.1177/0165025410367443

Kamii, C., Kirkland, L., & Lewis, B. A. (2001). Representation and abstraction in young children's numerical reasoning. In A. A. Cuoco & F. R. Curcio (Eds.), *The roles of representation in school mathematics: 2001 Yearbook of the National Council of Teachers of Mathematics* (pp. 24–34). National Council of Teachers of Mathematics.

Kamina, P., & Iyer, N. (2009). From concrete to abstract: Teaching for transfer of learning when using manipulatives. *NERA Conference Proceedings 2009*. https://digitalcommons.lib.uconn.edu/nera_2009/6/

Karpov, Y. V., & Haywood, H. C. (1998). Two ways to elaborate Vygotsky's concept of mediation. *American Psychologist, 53*(1), 27–36. https://doi.org/10.1037/0003-066X.53.1.27

Kato, Y., Kamii, C., Ozaki, K., & Nagahiro, M. (2002). Young children's representations of groups of objects: The relationship between abstraction and representation. *Journal for Research in Mathematics Education*, *33*(1), 30–45. https://doi.org/10.2307/749868

Kelley, D. (1984). A theory of abstraction. *Cognition & Brain Theory*, *7*(3–4), 329–357.

Kelley, D. (2010). *A theory of abstraction*. The Atlas Society.

Kesicioğlu, O. S., Alisinanoğlu, F., & Tuncer, A. T. (2011). The analysis of kindergarteners' recognition degrees of geometric shapes. *Elementary Education Online*, *10*(3), 1093–1111. https://dergipark.org.tr/tr/download/article-file/90637

Klein, A., Starkey, P., & Wakeley, A. (2004). Young children's composition of geometric figures: A learning trajectory. *Mathematical Thinking and Learning, 6*(2), 163-184.

Korkmaz, H. İ. (2017). Investigating kindergartners' geometric and spatial thinking skills: In context of gender and age. *European Journal of Education Studies*, 3(9), 55–69. https://doi.org/10.5281/zenodo.845497

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*(6), 2797–2822. https://doi.org/10.2307/1131753

Lehrer, R., Jenkins, M., & Osana, H. (1999). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments* *for developing understanding of geometry and space* (pp. 137–167). Lawrence Erlbaum Associates.

Lipovec, A. (2009). *Prototypical reasoning in developing early geometric concepts* [Conference paper]. 3rd International Conference Curriculums of the Early and Compulsory Education, Zadar, Croatia.

Liu, J., Golinkoff, R. M., & Sak, K. (2001). One cow does not an animal make: Young children can extend novel words at the superordinate level. *Child Development, 72*(6), 1674-1694. https://doi.org/10.1111/1467-8624.00372

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development, 2*(2), 189–219. https://doi.org/10.1207/S15327647JCD0202_4

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *AISHE-J: The All Ireland Journal of Teaching & Learning in Higher Education*, 8(3), 3351-33514.

Manning, J. P. (2005). Rediscovering Froebel: A call to re-examine his life & gifts. *Early Childhood Education Journal*, *32*, 371–376. https://doi.org/10.1007/s10643-005-0004-8

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M. S. C., & Westermann, G. (2007). *Neuroconstructivism: How the brain constructs cognition* (Vol. 1). Oxford University Press.

Martinez, S. F., & Huang, X. (2011). Epistemic groundings of abstraction and their cognitive dimension. *Philosophy of Science*, *78*(3), 490–511. https://doi.org/10.1086/660305

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding*. Shambhala Publications.

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. B., & Malin, A. (2008). Preschool influences on mathematics achievement. *Science*, *321*(5893), 1161–

1162. https://doi.org/10.1126/science.1158808

Millidge, B. (2021). Towards a mathematical theory of abstraction. *arXiv*. https://doi.org/10.48550/arXiv.2106.01826

Mistretta, R. M. (2000). Enhancing geometric reasoning. *Adolescence*, *35*(138), 365–379.

Mitchelmore, M. C., & White, P. (1995). Abstraction in mathematics: Conflict, resolution and application. *Mathematics Education Research Journal*, *7*(1), 50–68. https://doi.org/10.1007/BF03217275

Mitchelmore, M. C., & White, P. (2000). Development of angle concepts by progressive abstraction and generalisation. *Educational Studies in Mathematics*, *41*(3), 209–238. https://doi.org/10.1023/A:1003927811079

Mitchelmore, M., & White, P. (2007). Abstraction in mathematics learning. *Mathematics Education Research Journal*, *19*(2), 1–9. https://doi.org/10.1007/BF03217452

Mohr, D., Walcott, C., & Kastberg, S. (2007). The growth of the dynamic figural concept: Sense-making strategies applied to conceptions of shape. In D. K. Pugalee, A. Rogerson, & A. Schinck (Eds.), *Proceedings of the Ninth International Conference of Mathematics Education in a Global Community* (UNC Charlotte). UNC Charlotte.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*(1), 5-15. https://doi.org/10.1037/0096-3445.131.1.5

Nemirovsky, R., & Ferrara, F. (2009). Mathematical imagination and embodied cognition. *Educational Studies in Mathematics, 70*(2), 159-174. https://doi.org/10.1007/s10649-008-9150-4

Ness, D., & Farenga, S. J. (2007). *Knowledge under construction: The importance of play in developing children's spatial and geometric thinking*. Rowman & Littlefield.

Ness, D., & Farenga, S. J. (2016). Blocks, bricks, and planks: Relationships between affordance and visuo-spatial constructive play objects. *American Journal of Play*, *2*(8), 201–227. https://files.eric.ed.gov/fulltext/EJ1096889.pdf

Nokes, T. J., & Ohlsson, S. (2001). How is abstract, generative knowledge acquired? A comparison of three learning scenarios. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 710–715). Lawrence Erlbaum Associates.

Núñez, R. E. (2000). *Mathematical idea analysis: What embodied cognitive science can say about the human nature of mathematics.* In Proceedings of the Conference of the International Group for the Psychology of Mathematics Education (24th, Hiroshima, Japan, July 23-27), Vol. 1. ERIC.

Nurhasanah, F., Kusumah, Y. S., & Sabandar, J. (2017). Concept of triangle: Examples of mathematical abstraction in two different contexts. *International Journal on Emerging Mathematics Education*, *1*(1), 53–70. https://doi.org/10.12928/ijeme.v1i1.5782

Nurhasanah, F., Sabandar, J., & Kusumah, Y. S. (2013, March 17–22). *Abstraction processes in learning geometry using GSP* [Paper presentation]. 6th East Asia Regional Conference on Mathematics Education (EARCOME6), Phuket, Thailand. https://doi.org/10.13140/2.1.1452.0005

Ontario Ministry of Education. (2003). *Early math strategy: The report of the Expert Panel on Early Math in Ontario*. https://etfo-ots.ca/wp-content/uploads/2013/08/Early-Math-Strategy.pdf

Ontario Ministry of Education. (2019, August 28). *First year investment of Ontario's four-year math strategy announced* [Press release]. https://news.ontario.ca/edu/en/2019/08/first-year-investment-of-ontarios-four-year-math-strategy-announced.html

Oughton, R., Wheadon, D. M., Bolden, D. S., Nichols, K., Fearn, S., Darwin, S., Dixon-Jones, S., Mistry, M., Peyerimhoff, N., & Townsend, A. (2023). Going round in circles: Geometry in the early years. *Teaching Mathematics and its Applications, (286)*, 29–34. https://www.atm.org.uk/Mathematics-Teaching-Journal-Archive/177718

Ozerem, A. (2012). Misconceptions in geometry and suggested solutions for seventh grade students. *Procedia – Social and Behavioral Sciences*, *55*, 720–729. https://tinyurl.com/k9cs957z

Palmer, L. (1912). The relation between the ideal and the practical in the kindergarten programme. *The Elementary School Teacher*, *13*(6), 25-30. https://doi.org/10.1086/453778

Piaget, J. (2006). *The mechanisms of perception*. Routledge. (Original work published 1969)

Ponte, R., Viseu, F., Neto, T. B., & Aires, A. P. (2023). Revisiting manipulatives in the learning of geometric figures. *Frontiers in Education*, *8*, Article 1217680. https://doi.org/10.3389/feduc.2023.1217680

Posnansky, C. J., & Neumann, P. G. (1976). The abstraction of visual prototypes by children. *Journal of Experimental Child Psychology*, *21*(3), 367–379. https://doi.org/10.1016/0022-0965(76)90067-9

Pratt, D., & Noss, R. (2010). Designing for mathematical abstraction. *International Journal of Computers for Mathematical Learning*, *15,* 81–97. https://doi.org/10.1007/s10758-010-9160-z

Quinn, P. C. (1987). The categorical representation of visual pattern information by young infants. *Cognition, 27*(2), 145-179. https://doi.org/10.1016/0010-0277(87)90017-5

Reinhold, S., Downton, A., & Livy, S. (2017). *Revisiting Friedrich Froebel and his Gifts for*

Kindergarten: What are the benefits for primary mathematics education? (ED589542). ERIC. https://eric.ed.gov/?id=ED589542

Resnick, I., Verdine, B. N., Golinkoff, R., & Hirsh-Pasek, K. (2016). *Geometric toys in the attic? A corpus analysis of early exposure to geometric shapes. Early Childhood Research Quarterly, 36*, 358–365. https://doi.org/10.1016/j.ecresq.2016.01.007

Russell, B. (1897). An essay on the foundations of geometry. *Revue de Métaphysique et de Morale*, *6*(3), 354–380. https://philpapers.org/rec/RUSAEO-5

Sarama, J., & Clements, D. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. Routledge. https://doi.org/10.4324/9780203883785

Sarama, J., & DiBiase, A. M. (Eds.). (2004). *Engaging young children in mathematics standards for early childhood mathematics education.* Routledge.

Sarama, J., & Clements, D. H. (2009). Young children's composition of geometric figures: A learning trajectory. *Mathematical Thinking and Learning, 6*(2), 165–184. https://doi.org/10.1207/S15327833MTL0602_5

Sarama, J., & Clements, D. H. (2016). Physical and virtual manipulatives: What is "concrete"? In P. S. Moyer-Packenham (Ed.), *International perspectives on teaching and learning mathematics with virtual manipulatives* (Vol. 3, pp. 71–93). Springer. https://doi.org/10.1007/978-3-319-32718-1_4

Satlow, E., & Newcombe, N. (1998). When is a triangle not a triangle? Young children's developing concepts of geometric shapes. *Cognitive Development, 13*(4), 547–559. https://doi.org/10.1016/S0885-2014(98)90006-5

Scheiner, T., & Pinto, M. M. F. (2014). Cognitive processes underlying mathematical concept construction: The missing process of structural abstraction. In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA*

36 (Vol. 5, pp. 105–112). PME. https://files.eric.ed.gov/fulltext/ED600022.pdf

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 69*(Suppl 2), 107–131. https://doi.org/10.1007/s11577-017-0454-1

Sen, A. P., & Kotil, C. (2022). Re-evaluating Froebel Gifts: A tool that embodies three dimensions in space to enhance pre-school children's geometry skills. *International Journal of Educational Studies in Mathematics*, *9*(2), 129–145. https://doi.org/10.17278/ijesim.987450

Sharma, S. (2019a). Developing an understanding of shapes and their representations in multilingual primary school classes: A systematic literature review. In G. Hine, S. Blackley, & A. Cooke (Eds.), *Proceedings of the 42nd annual conference of the Mathematics Education Research Group of Australasia* (pp. 676–683). MERGA. https://files.eric.ed.gov/fulltext/ED604188.pdf

Sharma, S. (2019b). Use of theories and models in geometry education research: A critical review. *Waikato Journal of Education*, *24*(1), 43–54. https://doi.org/10.15663/wje.v24i1.644

Shipley, E. F., Kuhn, I. F., & Madden, E. C. (1983). Mothers' use of superordinate category terms. *Journal of Child Language, 10*(3), 571–588. https://doi.org/10.1017/S0305000900005377

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14. https://doi.org/10.3102/0013189X015002004

Silver, A. M., & Libertus, M. E. (2022). Environmental influences on mathematics performance in early childhood. *Nature Reviews Psychology*, *1*, 407–418. https://doi.org/10.1038/s44159-022-00061-z

Sinclair, N., Bartolini Bussi, M. G., de Villiers, M., Jones, K., Kortenkamp, U., Leung, A., & Owens, K. (2016). Recent research on geometry education: An ICME-13 survey team report. *ZDM*, *48*(5), 691–719. https://doi.org/10.1007/s11858-016-0796-6

Sinclair, N., & Bruce, C. D. (2015). New opportunities in geometry education at the primary school. *ZDM: The International Journal on Mathematics Education, 47*(3), 319-329. https://doi.org/10.1007/s11858-015-0693-4

Sinclair, N., & Moss, J. (2012). The more it changes, the more it becomes the same: The development of the routine of shape identification in dynamic geometry environment. *International Journal of Educational Research*, *51–52*, 28–44. https://doi.org/10.1016/j.ijer.2011.12.009

Spiegel, H., & Ginat, D. (2017). Decomposition considerations in geometry. In B. Kaur, W. K. Ho, T. L. Toh, & B. H. Choy (Eds.), *Proceedings of the 41st conference of the International Group for the Psychology of Mathematics Education (Vol. 4, pp. 209-216).* PME.

Sriraman, B., & English, L. D. (2005). On the teaching and learning of Dienes' principles. *Zentralblatt für Didaktik der Mathematik (ZDM), 37*(3), 258-262. https://doi.org/10.1007/s11858-005-0018-0

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19*(1), 99–120. https://doi.org/10.1016/j.ecresq.2004.01.002

Swafford, J. O., Jones, G. A., & Thornton, C. A. (1997). Increased knowledge in geometry and instructional practice. *Journal for Research in Mathematics Education, 28*(4), 467-483. https://doi.org/10.2307/749683

Teppo, A. (1991). Van Hiele levels of geometric thought revisited. *The Mathematics Teacher,* *84*(3), 210-221. Retrieved from https://www.jstor.org/stable/27967094

Tovey, H. (2016). *Bringing the Froebel approach to your early years practice*. Routledge.

Tovey, H. (2019). *Froebel's principles and practice today.* Froebel Trust. https://www.froebel.org.uk/uploads/documents/FT-Froebels-principles-and-practice-today.pdf

Usiskin, Z. P. (1972). The effects of teaching Euclidean geometry via transformations on student achievement and attitudes in tenth-grade geometry. *Journal for Research in Mathematics Education*, *3*(4), 249–259. https://doi.org/10.2307/748492

Verdine, B. N., Zimmermann, L., Foster, L., Marzouk, M. A., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. (2019). *Effects of geometric toy design on parent–child interactions and spatial language. Early Childhood Research Quarterly, 46*, 126–141. https://doi.org/10.1016/j.ecresq.2018.03.015

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2017). Spatial skills, their development, and their links to mathematics. *Monographs of the Society for Research in Child Development*, *82*(1), 7–30. https://doi.org/10.1111/mono.12280

Verdine, B. N., Lucca, K. R., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2016). The shape of things: The origin of young children's knowledge of the names and properties of geometric forms. *Journal of Cognition and Development*, *17*(1), 142–161. https://doi.org/10.1080/15248372.2015.1016610

Verdine, B. N., Bunger, A., Athanasopoulou, P., Golinkoff, R. M., & Hirsh-Pasek, K. (2017). *Shape up: An eye-tracking study of preschoolers' shape name processing and spatial development. Journal of Experimental Psychology: Human Perception and Performance.* https://doi.org/10.1037/dev0000384

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development, 85*(3), 1062–1076. https://doi.org/10.1111/cdev.12165

Vinner, S., & Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for Research in Mathematics Education, 20*(4), 356–366. https://doi.org/10.2307/749441

Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology*, *14*(3), 293–305. https://doi.org/10.1080/0020739830140305

Vinner, S., & Hershkowitz, R. (1980). Concept images and common cognitive paths in the development of some simple geometrical concepts. *Proceedings of the Fourth International Conference for the Psychology of Mathematics Education* (PME 4), 177–184. University of California, Berkeley.

von Marenholtz-Bülow, B. (2007). *How kindergarten came to America: Friedrich Froebel's radical vision of early childhood education*. The New Press.

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

Walcott, C., Mohr, D., & Kastberg, S. E. (2009). Making sense of shape: An analysis of children's written responses. *Journal of Mathematical Behavior, 28*(1), 30–40. https://doi.org/10.1016/j.jmathb.2009.04.001

Walker, C. M., Bridgers, S., & Gopnik, A. (2016). The early emergence and puzzling decline of relational reasoning: Effects of knowledge and search on inferring abstract concepts. *Cognition*, *156*, 30–40. https://doi.org/10.1016/j.cognition.2016.07.008

Walker, C. M., Hubachek, S. Q., & Vendetti, M. S. (2018). Achieving abstraction: Generating far analogies promotes relational reasoning in children. *Developmental Psychology*, *54*(10), 1833–1841. https://doi.org/10.1037/dev0000581

White, P., & Mitchelmore, M. C. (2010). Teaching for abstraction: A model. *Mathematical Thinking and Learning*, *12*(3), 205–226. https://doi.org/10.1080/10986061003717476

Wiles, P., & Anderson, R. (2019). Using a discursive framework to analyze geometric learning and instruction. In S. Otten, A. G. Cadela, Z. de Araujo, C. Haines, & C. Munter (Eds.), *Proceedings of the 41st Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Retrieved from ERIC: https://eric.ed.gov/?id=ED606944*

Willingham, D. T. (2009). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom*. Jossey-Bass.

Wilson, S. (1967). The "Gifts" of Friedrich Fröbel. *Journal of the Society of Architectural Historians, 26*(4), 238-241. https://doi.org/10.2307/988449

Wolfgang, C. H., Stannard, L. L., & Jones, I. (2001). Block play performance among preschoolers as a predictor of later school achievement in mathematics. *Journal of Research in Childhood Education, 15*(2), 173-180. https://doi.org/10.1080/02568540109594958

Woodard, C. (1979). *Gifts from the Father of the Kindergarten. The Elementary School Journal, 79*(3), 136-141. https://eric.ed.gov/?id=EJ201809

Yee, E. (2019). Abstraction and concepts: When, how, where, what and why? *Language, Cognition and Neuroscience*, *34*(10), 1257–1265. https://doi.org/10.1080/23273798.2019.1660797

Yeşil Dağlı, Ü., & Halat, E. (2016). Young children's conceptual understanding of triangle. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(2), 189–202. https://doi.org/10.12973/eurasia.2016.1398a

Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. Developmental Psychology, 53(6), 1098–1114. https://doi.org/10.1037/dev0000302

Zuckerman, O. (2006, January). *Historical overview and classification of traditional and digital learning objects*. MIT Media Laboratory.

Appendix A

Data Collection Tools, Consent Forms, and Invitation Letter

Geometric Shape Recognition Test (Aslan, 2007)

TEST REGISTRATION FORM

First name, Last name:
Date of birth:
Date:
Gender:
Time slot:
Day care:

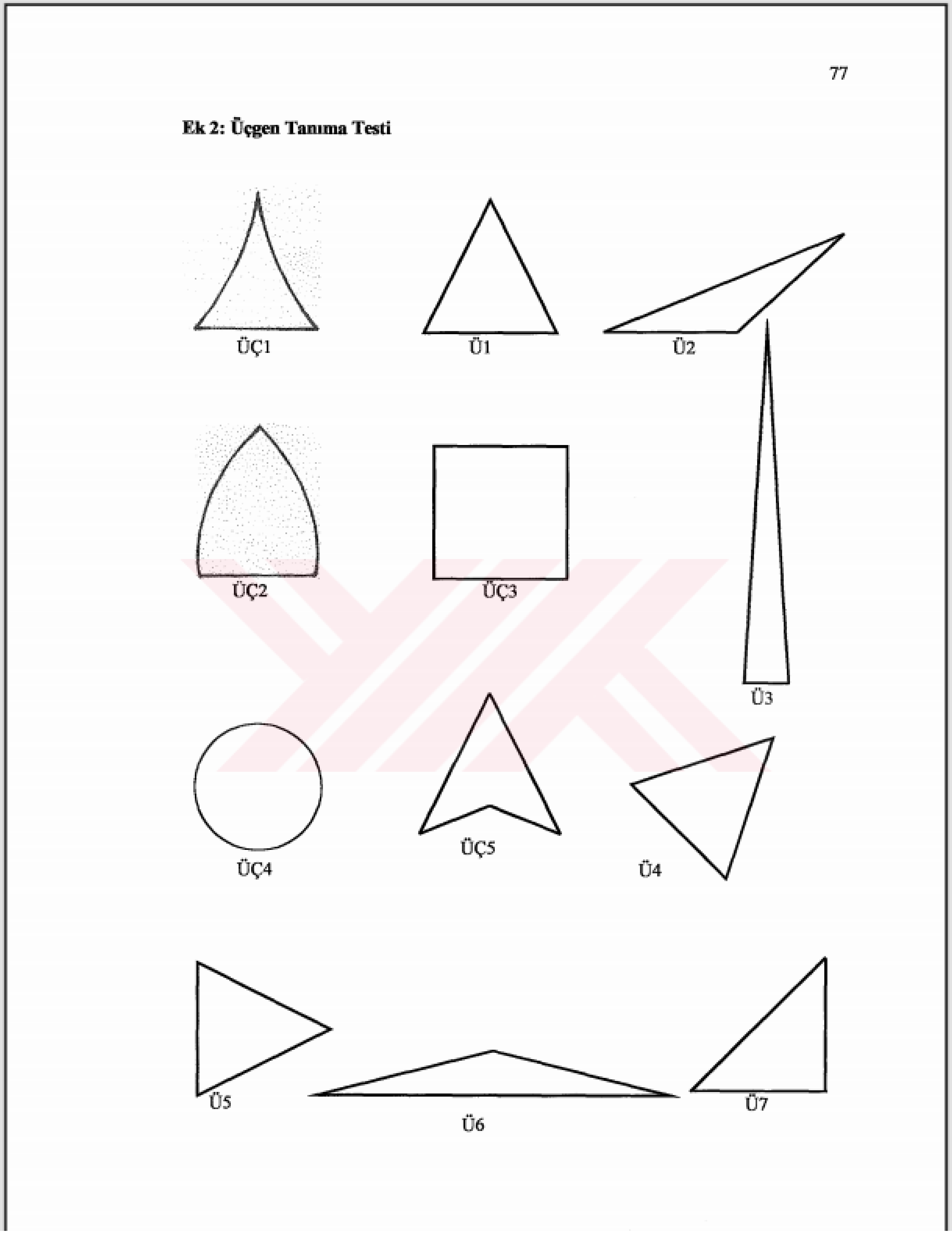
Code of the shape	**Recognition of the shape**	**Descriptions**
UC1		
U1		
U2		
UC2		
UC3		
U3		
UC4		
UC5		
U4		
U5		

U6		
U7		
DC1		
DC2		
D1		
DC3		
D2		
DC4		
D3		
DC5		
D4		
DC6		
DC7		
D5		
KC1		
KC2		
K1		

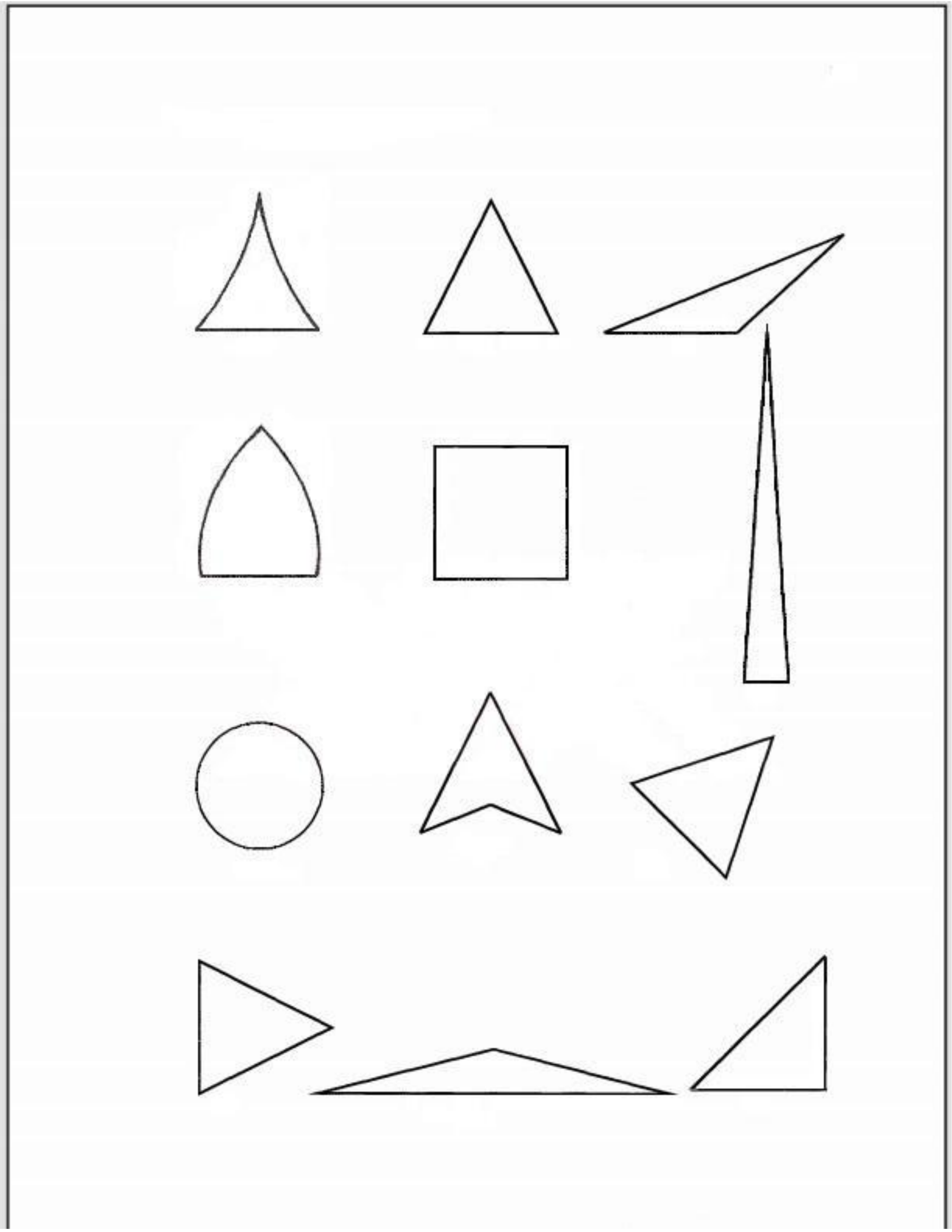
K2		
K3		
KC3		
KC4		
KC5		
K4		
KC6		
KC7		
KC8		
DAC1		
DAC2		
DA1		
DAC3		
DAC4		
DA2		
DA3		
DAC5		

DA4		
DAC6		
DAC7		
DA5		
TOTAL		

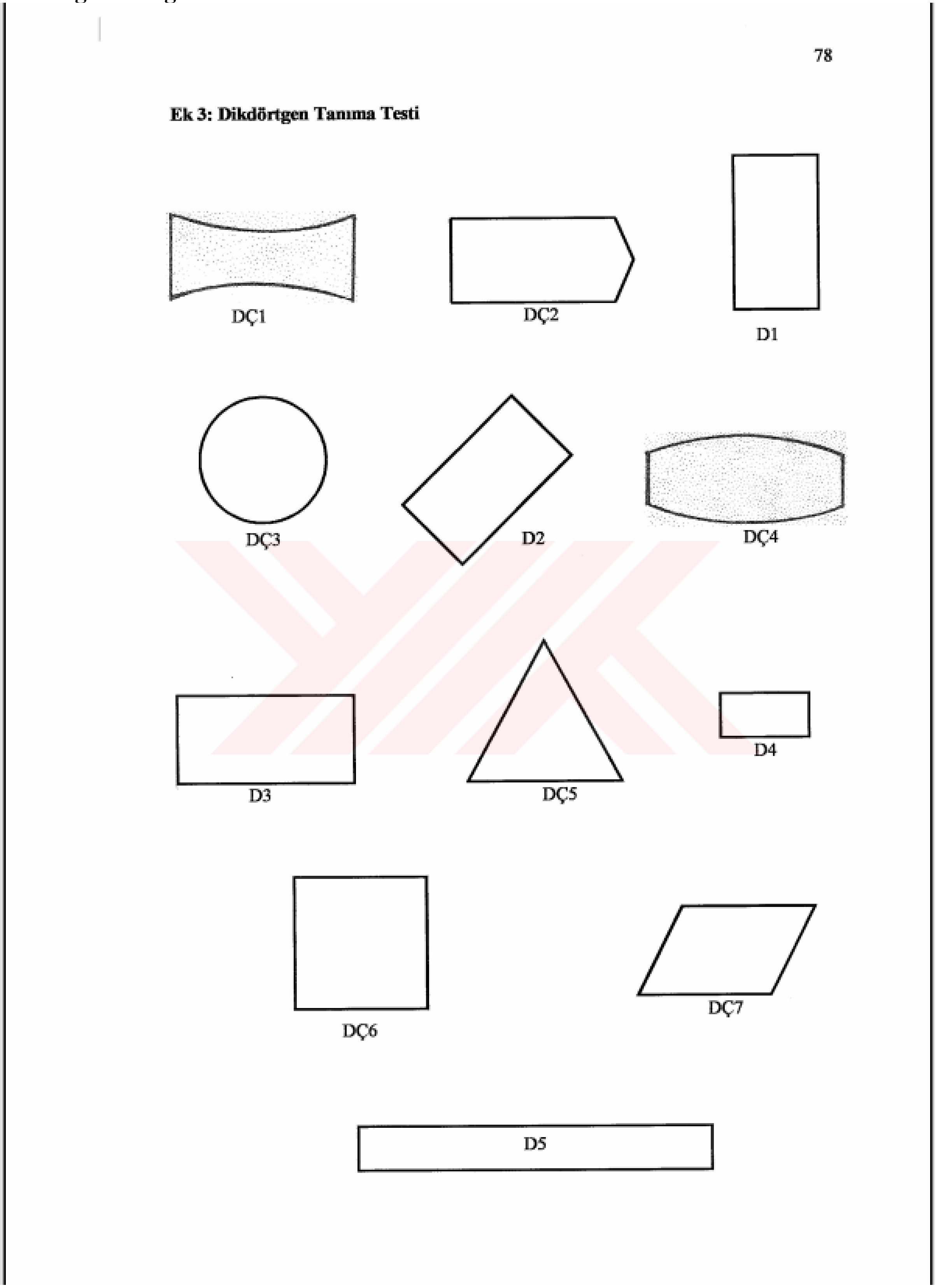
Triangle Recognition Task



Ek 2: Üçgen Tanıma Testi

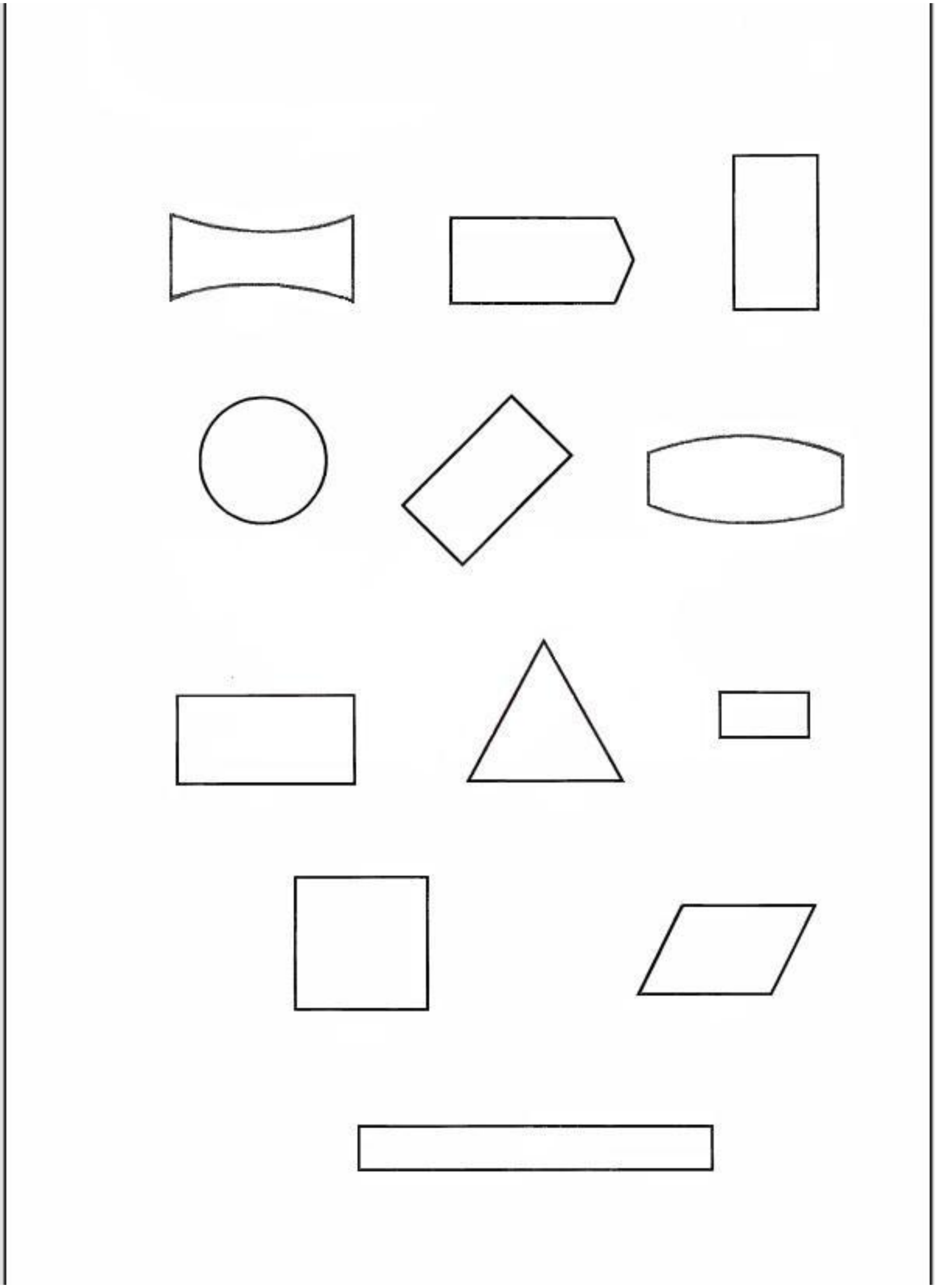


Rectangle Recognition Task



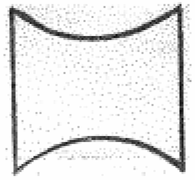
78

Ek 3: Dikdörtgen Tanıma Testi

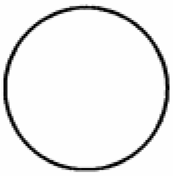


Square Recognition Task

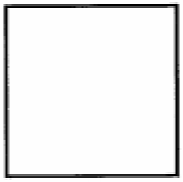
Ek 4: Kare Tanıma Testi



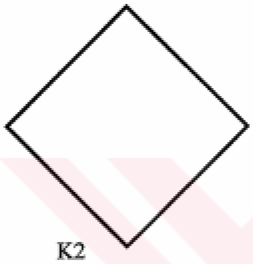
KÇ1



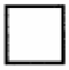
KÇ2



K1

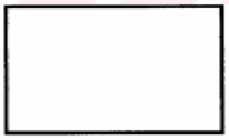


K2

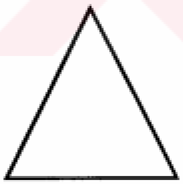


K3

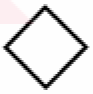
KÇ3



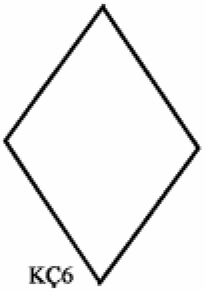
KÇ4



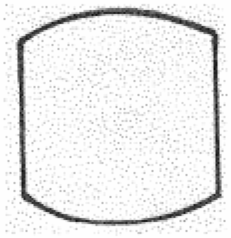
KÇ5



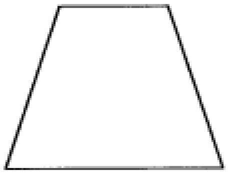
K4



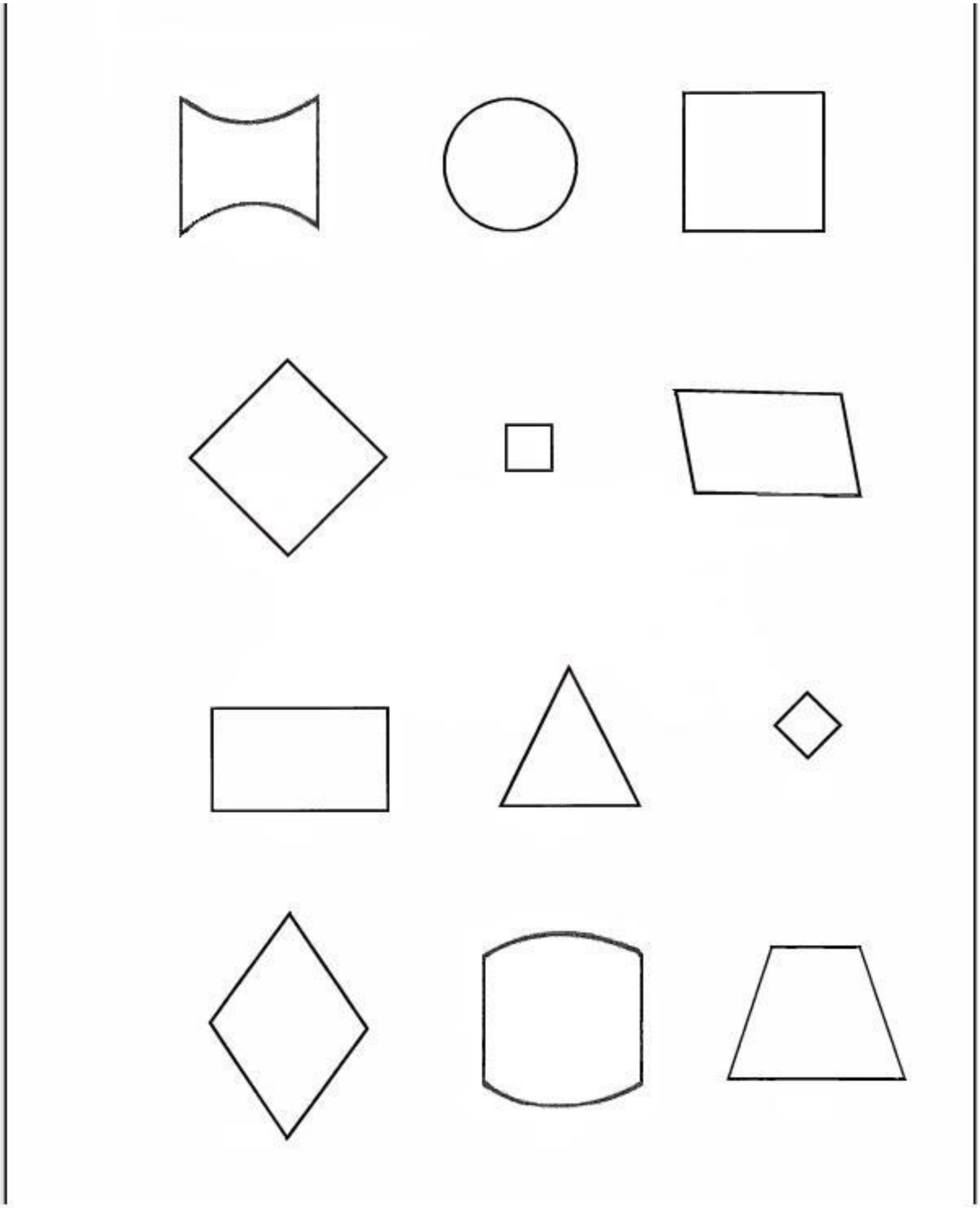
KÇ6



KÇ7

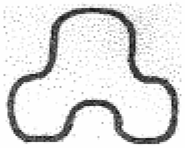


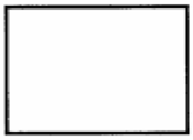
KÇ8

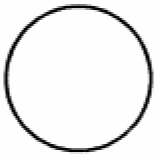


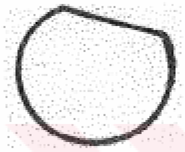
Circle Recognition Task

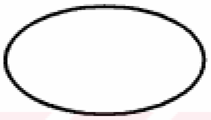
Ek 6: Daire Tanıma Testi

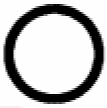

DAÇ1


DAÇ2

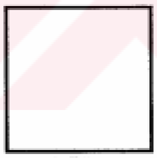

DA1

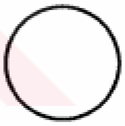

DAÇ3

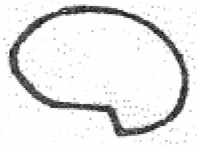

DAÇ4

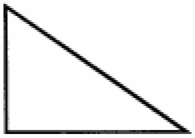

DA2


DA3

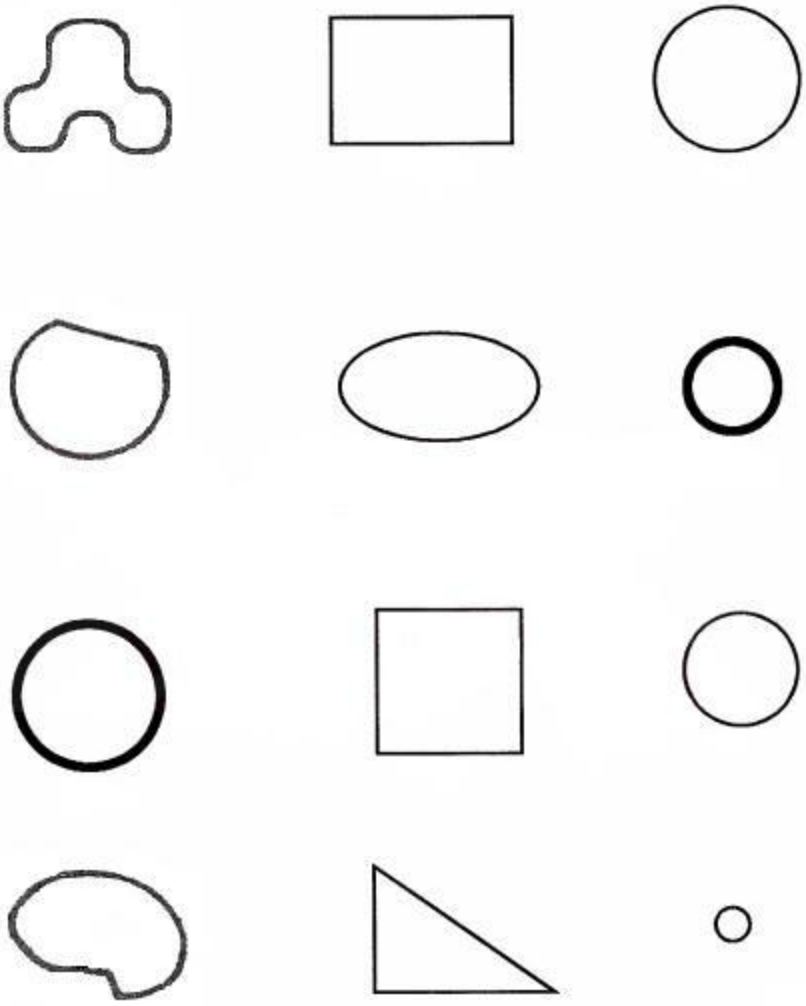

DAÇ5


DA4


DAÇ6


DAÇ7


DA5



Semi-structured Interview Questions for Teachers

Abstract geometry concepts/Interview Themes	**Activities/Instructions**	**Vocabulary**	**Manipulatives**
Two dimensional shapes, planes	1.1. How was your experience of helping the children build the concept of planes or surfaces of three-dimensional shapes, or in other words two-dimensional shapes? 1.2. What kind of activities did you do with the children about and/or including the concept of planes (surfaces of two- and three-dimensional shapes)? 1.3. Did you or the children make or indicate any connection between the concept of planes and any other geometrical or mathematical concept such as edges, vertices, area, angle, etc?	1.4. What words did you or the children use in the classroom for this concept such as surface, top, face, etc? 1.5. What words did you use or the children use for: square 1.6. rectangle 1.7. rhombus 1.8. paralellogram 1.9. trapezoid 1.10. triangle 1.11. isosceles triangle 1.12. equilateral triangle 1.13. right triangle 1.14. obtuse triangle 1.15. circle 1.16. Did you guide children to use appropriate vocabulary when they use another word to refer to any of the shapes above? 1.17. Did you explain the children why they must use specific terms for specific shapes? If so, how	1.18. Did you use any manipulatives in your activities? 1.19. How did you incorporated the manipulative(s) in any of your activities to facilitate their learning about the concept of planes?

		did you explain it? Can you provide an example from the class experiences?	
Three dimensional shapes, solids	2.1. How was your experience of helping the children build the concept of solids or three-dimensional shapes? 2.2. What kind of activities did you do with the children about and/or including the concept of solids (three-dimensional shapes)? 2.3. Did you or the children make or indicate any connection between the concept of solids and any other geometrical or mathematical concept such as planes, edges, vertices, area, angle, etc?	2.4. What words did you or the children use in the classroom for the concept of solids (e.g. three dimensional shapes, solids, objects)? 2.5. What words did you use or the children use for: cube 2.6. sphere 2.7. rectangular 2.8. prism 2.9. triangular prism 2.10. cylinder 2.11. Did you guide children to use appropriate vocabulary when they used another word to refer to any of the shapes above? 2.12. Did you explain to he children why they must use specific terms for specific shapes? If so, how did you explain it? Can you provide an example from the class experiences?	2.13. Did you use any manipulatives in your activities? 2.14. How did you incorporated the manipulative(s) in any of your activities to facilitate their learning about the concept of solids?

Side, line, edge	3.1. How was your experience of guiding the children build the concept of sides or edges of two- and three-dimensional shapes? 3.2. What kind of activities did you do with the children about and/or including the concept of sides/edges of surfaces and three-dimensional shapes? 3.3. Did the children make or indicate any connection between the concept of side/edge/line and any other geometrical or mathematical concept such as surface, angle, vertices, etc?	3.4. What words did you or the children use in the classroom for this concept such as line, edge, side? 3.5. Did you guided children's to use appropriate vocabulary when they use another word to refer side, line, or edges of shapes? 3.6. Did you explain the children why they must use that specific term? If so, how did you explain it?	3.7. Did you use any manipulatives in your activities? 3.8. How did you incorporated the manipulative(s) in any of your activities to facilitate their learning about the concept of sides/edges in two- and three-dimensional geometric shapes?
Vertices	4.1. How was your experience of helping the children build the concept of vertices of two- and three-dimensional shapes? 4.2. What kind of activities did you do with the children about and/or including the concept of vertices of two- and three-dimensional shapes? 4.3. Did you or the children make or indicate any	4.4. What words did you or the children use in the classroom for this concept such as point, corner, etc? 4.5. Did you guided children's to use appropriate vocabulary when they use another word to refer vertices of shapes? 4.6. Did you explain the children why they must use that specific	4.7. Did you use any manipulatives in your activities? 4.8. How did you incorporated the manipulative(s) in any of your activities to facilitate their learning about the concept of vertices in two- and three-dimensional geometric shapes?

	connection between the concept of vertices and any other geometrical or mathematical concept such as edges, lines, area, angle, etc?	term? If so, how did you explain it?	
Angles	5.1. How was your experience of helping the children build the concept of angles of two- and three-dimensional shapes? 5.2. What kind of activities did you do with the children about and/or including the concept of angles of two- and three-dimensional shapes? 5.3. Did you or the children make or indicate any connection between the concept of angles and any other geometrical or mathematical concept such as edges, vertices, area, width, etc?	5.4. What words did you or the children use in the classroom for this concept such as corner, width? 5.5. How did you guided children to use appropriate vocabulary when they use another word to refer angles? 5.6. Did you explain the children why they must use that specific term? If so, how did you explain it?	5.7. Did you use any manipulatives in your activities? 5.8. How did you incorporated the manipulative(s) in any of your activities to facilitate their learning about the concept of vertices in two- and three-dimensional geometric shapes?

Observation Focus Form

Observation Focus	**Observation points for the teachers**	**Observation points for the children**	**Date Time Round 1 / 2**
Activities/ Instructions	-What kinds of activities do the teachers do with the students to investigate geometric objects and shapes? -How do the teachers guide/lead/instruct children to compose or decompose shapes? -What kinds of activities do the teachers do with children regarding representing/translating geometry concepts into general knowledge? -What kinds of activities do the teachers do with children to inquire about similarities between geometric objects and shapes?	-What kind of activities do the children choose to do? -What kind of questions do the children ask? -What connections do children make between geometry concepts and other knowledge? -Do the children mention anything about the similarities and differences between the shapes? -Do the children compose and decompose two- and three-dimensional shapes?	
Vocabulary	-Do teachers use the terminology “square”, “rectangle”, “triangle”, “equilateral triangle”, “isosceles triangle”, “scalene triangle”, “right triangle”, “circle”, “rhombus”, “trapezoid”, “parallelogram”, “pentagon”, “hexagon” when introducing or describing a two-dimensional shape? -Do teachers use the terminology “angle”, “plane”, “line”, “side”,	-What kind of conversations do the children have with their peers or with their teachers? -What words do the children use to describe a square?	

	"edge", "vertices", "area" when describing a shape? -Do teachers use the terminology "cube", "rectangular prism", "triangular prism", "cylinder" when introducing or describing a three-dimensional shape? -How do teachers guide the children to use the abstract concepts (e.g. angle, line, plane, etc.) in appropriate contexts during activities? -How do the teachers guide the children to use the correct terminology when necessary?	-What words do the children use to describe a triangle? -What words do the children use to describe a circle? -What words do the children use to describe a rectangle? -How do children pronounce the names of the shapes they encounter with? -Do children distinguish appropriate words for two- and three-dimensional shapes? (e.g. cube, square, rectangle, rectangular prism) -What words do children use to describe components of the two- and three-dimensional shapes such as edges, vertices, planes, surfaces, sides, corners, etc.?	
Manipulatives	-Do teachers employ manipulatives to introduce or describe abstract concepts in geometry? If so, what kind of materials are they using?	-In which conditions and situations do children tend to use manipulatives?	

	-Do teachers provide manipulatives for children to inquire about a shape? If so, what kind of materials are they using? -Do teachers provide the children with appropriate manipulatives that represent the components of shapes? If so, what kind of materials are they using? -Do teachers provide opportunities to children to compose and decompose shapes with the manipulatives?	-What kind of manipulatives do the children use during activities? -In which ways do they use the manipulatives that they are using? -Do children use specific manipulatives to represent two-dimensional shapes/planes? -Do children use specific manipulatives to represent edges/sides of two- and three-dimensional shapes? -Do children use specific manipulatives to represent vertices?	

Parental Consent Form

Parental Consent Form

Title of Study: The Role of Composition, De-composition, and Comparison Activities in 3 – 4 Years Old Children's Abstraction Processes in Shape Recognition

Principal Student Investigator: Ayse Pinar Sen, Ph.D. Candidate, Faculty of Education, Brock University,Ps17ob@brocku.ca

Principal Investigator: Dr. Mary Louise Vanderlee, Professor, Faculty of Education, Brock University, mvanderlee@brocku.ca

Introduction and Purpose of the Study

The purpose of this form is to provide you with information that may affect your decision as to whether to allow your child to participate in this research study. I have developed a new math program that will support a child's ability to learn about and recognize shapes. This learning is associated with abstract thinking which will be assessed along with shape recognition to identify whether or not the math program developed is effective. Classroom observations will be conducted to understand how children learn about shapes and identify effective learning activities and materials. Your child will not be audio/video recorded.

What is my child going to be asked to do?

If you permit your child's participation in the study, I will also ask for their individual assent. In case your child loses interest, I will try to gain their attention/assent on the same day, with a maximum of two assent approaches to avoid undue pressure. With your consent, your child will be assigned to either the experimental or control group. Participants, with both consent and assent, will be asked shape-related questions at the study's beginning and end. In the experimental group, children will play with Froebel Gifts for 8 weeks during classroom activities, sitting at a designated table. The control group will continue regular activities in their classroom. If you choose not to participate, your child will engage with regular materials for 8 weeks in a separate group with one of their teachers during their regular schedule. Participating and non-participating children will remain in the classroom. Non-participating children will not be observed or participate in the pre- or post-test.

What are the risks involved in this study?

There are no foreseeable risks to participating in this study.

What are the possible benefits of this study?

Participation in the study is expected to benefit children by enhancing their abstract thinking skills in geometric shape recognition. Children will engage in activities from the Ayse Pinar Sen Geometry Curriculum (APSGC) and play with Froebel Gifts under teacher supervision. The application aims to foster the acquisition of abstract geometry concepts, increase mathematical and geometry vocabulary, and improve shape recognition skills. To address potential disparities between experimental and control classrooms, teachers in the control groups will receive APSGC and 10 sets of Froebel Gifts for each classroom after data collection, ensuring all children have the opportunity to develop abstract thinking skills in geometry.

Confidentiality

All information collected in this study will be confidential, with no personal identifying details used in reports. Records will be securely stored, with access restricted to the researcher and supervisor. Children's answers will be pseudonymized, and the data collected from all the children will be analyzed all together, without personal identification ensuring confidentiality. After the Ph.D. dissertation defense, hard copies will be shredded, and electronic files deleted. Participation is voluntary, and withdrawal is allowed during the recruitment process and before the pre-test of children. However, participants can no longer withdraw when the study is at advanced stages of data analysis or after information about the data is published in any form, (i.e., via conference, journal, or dissertation). To withdraw, send a request to the provided email of the researcher, resulting in permanent data deletion. If you decide to allow your child to participate in this research, please sign and send this consent form to me at ps17ob@brocku.ca prior to the observation procedure before January 22nd, 2024. Any questions can be directed to the Student

Investigator. Concerns about the research can be addressed to the Office of Research Ethics at Brock University at (905) 688-5550 Ext 3035, or via email at reb@brocku.ca. Thank you for your interest in the project.

Consent

I give permission for researchers to use a pseudonym to represent my child in this research. I acknowledge that no personal information about my child will be disclosed individually in any publications based on the research data. I agree to allow my child(ren) to participate in the described study, having made this decision after reading the Informed-Consent Letter. I am aware of the opportunity to seek additional details and can ask questions in the future.

____

Printed Full Name of Child (to be replaced with a pseudonym by the researcher)

____________________________________ ____________________

Signature of Parent(s) or Legal Guardian Date

____________________________________ ____________________

Signature of Investigator Date

Teacher Consent Form

Teacher Consent Form

Title of Study: The Role of Composition, De-composition, and Comparison Activities in 3 – 4 Years Old Children’s Abstraction Processes in Shape Recognition

Principal Student Investigator: Ayse Pinar Sen, Ph.D. Candidate, Faculty of Education, Brock University, Ps17ob@brocku.ca

Principal Investigator: Dr. Mary Louise Vanderlee, Professor, Faculty of Education, Brock University mvanderlee@brocku.ca

Introduction to my study: What is my purpose?

The purpose of this form is to provide you with information that may affect your decision as to whether to participate with your classroom in this research study. I, the researcher, have developed a new math program aimed at helping children learn and recognize shapes. This learning is associated with abstract thinking which will be assessed along with shape recognition to identify whether or not the math program developed is effective. Classroom observations will be conducted to understand how children learn about shapes and identify effective learning activities and materials. You and the children in your classroom will not be audio/video recorded. The procedures in this study, which will take place between …….. 2024 and ……. 2024, include the following expectations for you as a teacher.

What's Involved?

The study involves inviting classrooms to participate in an 8-week research project focusing on the recognition of geometric shapes among 3-4-year-old children. Teachers are asked to send the parental consent forms to parents for participation. The experimental group will engage in activities with Froebel Gifts within a new curriculum, while the control group will continue regular activities with standard materials. Participating and non-participating children will remain in the classroom. Non-participating children will not be observed or participate in the pre- or post-test. Teachers are asked to provide a separate table for participating children during the study. Face-to-face assessments will be conducted at the beginning and end of the study, lasting 25-30 minutes for each child. Observations of teacher-child discourse will occur for 3 hours, 2 days a week, over 10 weeks. Froebel Gifts and curriculum materials will be provided to the experimental group. Online training sessions for experimental group teachers will be scheduled. Face-to-face or online interviews with teachers will take place after post-test applications, lasting 45 minutes to an hour, with interview transcripts provided for verification within two weeks.

What are the risks involved in this study?

There are no foreseeable risks to participating in this study.

What are the possible benefits of this study?

Participation in the study offers potential benefits, including the professional development of teachers in early-years geometry education and practical insights into using Froebel Gifts for math and geometry activities. The APSG curriculum activities aim to enhance children's understanding of abstract geometry concepts, boost math and geometry vocabulary, and foster shape recognition skills, laying a solid foundation for future academic pursuits. To address potential disparities in abstract thinking skills between experimental and control classrooms, the Ayse Pinar Sen Geometry Curriculum (APSGC) and Froebel Gifts will be provided to control group teachers after data collection. This ensures all children have the opportunity to learn geometry through APSGC. As compensation, each participating classroom will receive 1 set of Froebel Gifts, even if they withdraw from the study after data collection.

Confidentiality

This consent form is provided because the daycare center's administrator has opted for the center's participation in this research study. However, the decision to participate is entirely voluntary, and the individual's choice will not be disclosed to the administrator. Participants can withdraw at any point during the recruitment process and before the pre-test and interview, without providing a reason. To withdraw, an email request should be sent to ps17ob@brocku.ca, resulting in the permanent deletion of all collected data. The study assures confidentiality, with information securely stored and accessible only to the researcher (Ayse Pinar Sen) and the Ph.D. supervisor (Dr. Mary Louise Vanderlee - Principal Investigator). Personal identifiers will be replaced with pseudonyms during analysis. After the Ph.D. dissertation defense, all hard copies and electronic files will be securely disposed of. To participate, individuals are invited to sign and send the consent form with any questions via email to ps17ob@brocku.ca. Thank you for your kind interest in this project.

Consent

I consent to the use of a pseudonym for myself and the children in my classroom in this research. I am aware that no personal information about me or the children will be disclosed in any publications based on the research data. I understand the option to withdraw from the study before the data collection phase concludes by revoking this consent. Having read the Informed-Consent Letter, I have had the chance to obtain additional details about the study and know that I can ask questions in the future. I agree to participate and allow the children in my classroom to take part in the described study.

____
Teacher's full name

______________________________ ________________
Signature of the teacher Date

______________________________ ________________
Signature of Investigator Date

Assent Script for Minors

Assent Script for Minors

Title of Study: The Role of Composition, De-composition, and Comparison Activities in 3 – 4 Years Old Children's Abstraction Processes in Shape Recognition

Invitation to participate in the research

I will read the following to invite each child to engage in an activity about shapes which is the Geometric Shape Recognition Test (Aslan, 2007), through the conversation script below.

"Hello (Insert Child's Name),

My name is Pinar. How are you?

I am very happy to meet you. I am a teacher and I have worked with other young children. I would like to know if you would like to join me at the table (point to table) for an activity. I would like to talk about some shapes like circles, squares, triangles, and rectangles with you. If you don't want to join me at the table, you can stay where you are. You don't have to follow me. I am going to walk over to the activity table, if you follow me, I will know that you would like to do the shape activities. Would you like to come with me to the table to do an activity with me or stay where you are?"

I will listen to the child's answer. If she/he says no, or indicates that she/he does not want to join the activity that I am inviting her/him to, I will thank her/him for their response and tell them that they can join their teachers and other friends in the classroom. If the child indicates that they would like to come with me to join the activity, then I will walk toward the table and if the child follows, and sits with me, then I will ask:

"Are you ready to begin?"

I will listen to their responses, and I will state:

"If you want to stop this activity, you can tell me to stop anytime you want."

If the child loses interest in the study (e.g., walks away to do something else), I will try to gain their attention/assent again later within the same day after an hour of my initial approach. In this case, I will approach the child for assent only one more time to mitigate any potential pressure on the child.

Invitation Letter for Parents

Invitation Letter for Parents

Title: The Role of Composition, De-composition, and Comparison Activities in 3 – 4 Years Old Children's Abstraction Processes in Shape Recognition

Principal Student Investigator: Ayse Pinar Sen, Ph.D. Candidate, Faculty of Education, Brock University

Principal Investigator: Dr. Mary Louise Vanderlee, Professor, Interim Dean of the Faculty of Education, Brock University

My name is Ayse Pinar Sen, and I am a Ph.D. candidate and an instructor at the Faculty of Education, at Brock University. My research is supervised by Professor Mary-Louise Vanderlee, Ed.D. I am currently recruiting for a study on the development of young children's abstract thinking when working with or playing with shapes. We know that the children's abstract thinking develops within their interactions with the teachers, with each other, and with the materials, vocabulary, activities, and instructions. However, we don't know which materials, instructions, activities, or discourse are more effective in developing children's abstract thinking skills in geometry. Therefore, this study focuses on the impact of an innovative curriculum on children's abstract thinking skills in geometry. Therefore, I need two groups:

- One to use the new curriculum (Ayse Pinar Sen Geometry Curriculum)
- The other is to continue with their regular programming.

This will allow me to evaluate the impact of the new curriculum.

I would like to invite your child(ren) between the ages of 3-4 to participate in this project. If you allow your child(ren) to participate in this research, please scan the QR code below for the parental consent form. I am also sharing the QR code for the assent script for minors to inform you regarding the assent process for your child.

QR code for Parental Consent Form

QR code for Assent Script for Minors

If you are interested in participating in this study, please kindly sign the consent form and send it to me, Ayse Pinar Sen via email at ps17ob@brocku.ca. Also, if you have any questions or want further information, please contact Ayse Pinar Sen, Ph.D. Candidate, Faculty of Education, Brock University via email: ps17ob@brocku.ca

This study has been reviewed and received ethics clearance through Brock University's Research

Ethics Board.

If you have any questions or concerns about your rights as a research participant, please contact the Brock University Research Ethics Officer at (905) 688-5550 Ext 3035, reb@brocku.ca.

Thank you,

Ayse Pinar Sen

Invitation Letter for Teachers

Invitation Letter for Teachers

Title: The Role of Composition, De-composition, and Comparison Activities in 3 – 4 Years Old Children's Abstraction Processes in Shape Recognition

Principal Student Investigator: Ayse Pinar Sen, Ph.D. Candidate, Faculty of Education, Brock University

Principal Investigator: Dr. Mary Louise Vanderlee, Professor, Interim Dean of the Faculty of Education, Brock University

My name is Ayse Pinar Sen, and I am a Ph.D. candidate and an instructor at the Faculty of Education, at Brock University. My research is supervised by Professor Mary-Louise Vanderlee, Ed.D. I am currently recruiting for a study on the development of young children's abstract thinking when working with or playing with shapes. We know that the children's abstract thinking develops within their interactions with the teachers, with each other, and with the materials, vocabulary, activities, and instructions. However, we don’t know which materials, instructions, activities, or discourse are more effective in developing children’s abstract thinking skills in geometry. Therefore, this study focuses on the impact of an innovative curriculum on children’s abstract thinking skills in geometry. Therefore, I need two groups:

- One to use the new curriculum (Ayse Pinar Sen Geometry Curriculum)
- The other is to continue with their regular programming.

This will allow me to evaluate the impact of the new curriculum.

I would like to invite you and the children in your classroom who are between the ages of 3-4 to participate in this project entitled *The Role of Composition, De-composition, and Comparison Activities in 3 – 4 Years Old Children’s Abstraction Processes in Shape Recognition.* In appreciation of your time and contributions as a participant in this study, we will compensate your classroom with a set of Froebel Gifts.

What are the procedures?

The procedures in this study, which will take place between …….. 2024 and …….

2024, include the following expectations for you as a teacher. You and your classroom can be assigned either as a control group or an experimental group.

If your classroom is assigned as a control group, you will be expected to:	If your classroom is assigned as an experimental group, you will be expected to:

● Continue implementing your regular activities. ● Host the researcher in your classroom for observation, 3 hours a day, 2 days a week, for 10 weeks. ● Answer the interview questions at the end of the research	● Participate in the Ayse Pinar Sen Geometry Curriculum training (two online sessions) ● Implement the Ayse Pinar Sen Geometry Curriculum 3 days a week, 2 hours a day, for 10 weeks ● Host the researcher in your classroom for observation, 3 hours a day, 2 days a week, for 10 weeks. ● Answer the interview questions at the end of the research

What will your students be asked to do?

I will ask all the children who consented to participate in this study, regardless of whether they are in the experimental or control group to answer the researcher's questions about the recognition of geometric shapes through the Geometric Shape Recognition Test, once at the beginning of the study and once at the end of the study. If the classroom of the child is assigned as the experimental group, the children will be invited to play with Froebel Gifts for 8 weeks when the teacher employs the manipulatives in the activities. If the classroom of the child is assigned as the control group, the teachers and the children will continue their regular activities with their regular manipulatives.

NOTE: You and your students **will not be** audio/video recorded during the observations and the intervention.

What's Involved?

If you consent to participate in my study with your classroom, I will send parental consent forms to be sent to parents. Based on parental consent, I will invite each child to participate in this research. This will provide each child with the ability to provide individual assent. If the children lose interest in the study (e.g., walk away to do something else), I will try to gain their attention/assent again later within the same day after an hour of my initial approach. In this case, I will approach the child for assent only one more time to mitigate any potential pressure on them.

Children who are participating will remain in the classroom. They will be invited to take a seat at a table on the edge of the classroom. If some of the parents in your classroom do not allow their children to participate in this research, then the early childhood educator or the assistant early childhood educator in your classroom will be asked to continue the regular activities with the regular materials and toys in your classroom with the children who have not assented, and whose parent's have not provided consent for their participation, as a separate group in the classroom. This group will not be observed or participate in the pre- or post-test.

After parents' consent is obtained from the volunteering parents, I will assign the classrooms as control and experimental groups. At the beginning of the research study, I will observe all the classrooms to take notes regarding the discourse between you and the children during the activities. I will only use a pen and paper to take notes. Following the first round of observations, I am going to ask you for your availability to make a schedule for the pre-test. Through the Geometric Shape Recognition Test (GSRT), I will ask each child who has assented, and whose parents have provided consent for their participation questions regarding the geometric shapes for 25-30 minutes in a face-to-face setting, in a quiet room at the daycare center. Following the pre-test applications, I will deliver 10 sets of Froebel Gifts and 2 copies of Ayse Pinar Sen

Geometry Curriculum (APSGC) to each experimental group. The children who have assented, and whose parents have provided consent for their participation in the experimental classrooms will be asked to play with Froebel Gifts when their teacher implements the APSGC activities for 8 weeks in a separate group, while the control classrooms use their regular manipulatives in the classroom. After 8 weeks, I will ask children the same questions in the pre-test, as a post-test. I will interview the teachers once I complete the post-test applications with the children. The interview procedure will take place face-to-face at the daycare centers where you are working or online through MS Teams, according to your preference and availability. Each interview will take 45 minutes to an hour. You don't have to worry about knowing the interview procedure, I will guide you through the entire interview session. I will send you the interview questions prior to the interview.

If you would like to participate in this research, please distribute the QR code for the parental consent form and the invitation letter for parents to the parents of 3-4-year-old children in your classroom. The following QR codes are linked to the teacher consent forms, consent forms for parents, and invitation letters for parents.

QR code for Teacher Consent Form

QR code for Parental Consent Form

QR code for the invitation letter for parents

If you are interested in participating in this study, kindly sign the consent form and send it via email at ps17ob@brocku.ca. Also, if you have any questions or want further information, please contact Ayse Pinar Sen, Ph.D. Candidate, Faculty of Education, Brock University via email: ps17ob@brocku.ca

This study has been reviewed and received ethics clearance through Brock University's Research Ethics Board (file #23-092 - VANDERLEE).

If you have any questions or concerns about your rights as a research participant, please contact the Brock University Research Ethics Officer at (905) 688-5550 Ext 3035, reb@brocku.ca.

Thank you,

Ayse Pinar Sen

Appendix B

Supporting Analyses for Main Quantitative Findings

Table of Contents

Section 4: ANCOVA and Post Hoc Analyses

Section 5: ANCOVA Model Diagnostics

Section 6: ANCOVA Summary and Residual Diagnostics

Section 7: Verbal Response Analyses (Chi-Square & Mann–Whitney U)

Section 1: Descriptive Statistics

Table 1a.

Descriptive Statistics for Pretest and Post-test Scores

Statistics	Total Score of Pretest	Total Score of Post-test
N		
Valid	93	91
Missing	0	2
Mean	37.80	38.86
Median	38.00	39.00
Std. Deviation	3.312	3.418
Skewness	-.475	-.236
Std. Error of Skewness	.250	.253
Kurtosis	.464	-.212
Std. Error of Kurtosis	.495	.500
Range	18	16
Minimum	27	31
Maximum	45	47

This table summarizes overall descriptive statistics (mean, median, standard deviation, skewness, kurtosis) for pretest and post-test scores. It provides a descriptive statistical summary, not an inferential test. This table includes all participants. The mean pretest score (M = 37.80, SD = 3.31) is slightly lower than the mean post-test score (M = 38.86, SD = 3.42), indicating modest improvement. Skewness and kurtosis values suggest near-normal distributions for both pretest (-0.475, 0.464) and post-test (-0.236, -0.212) scores. The range for post-test scores (31–47) is slightly narrower than the pretest range (27–45), reflecting greater consistency in post-test performance. The pretest skewness of -0.475 indicates a slight left skew, with more participants scoring in the higher range, while the posttest skewness of -0.236 suggests a distribution closer to symmetry. The kurtosis values of 0.464 (pretest) and -0.212 (posttest) indicate that both distributions are slightly flatter than a normal distribution. Overall, these descriptive statistics highlight modest improvement from pretest to posttest and confirm that participants' scores are concentrated in the upper range for both tests.

Table 1b.

Descriptive Statistics for Pretest and Post-test Scores by Group

Groups	Measure	N	Mean	Std. Deviation	Std. Error	95% CI Lower Bound	95% CI Upper Bound	Min.	Max.
Control Group (8 weeks)	Pretest Scores	27	37.81	2.95	0.57	36.65	38.98	33	43
	Posttest Scores	27	38.11	2.74	0.53	37.03	39.19	34	43
Experimental Group (8 weeks)	Pretest Scores	29	37.83	3.33	0.62	36.56	39.09	27	45
	Posttest Scores	29	40.31	3.53	0.66	38.97	41.65	31	47
Control Group (4 weeks)	Pretest Scores	20	38.45	3.98	0.89	36.59	40.31	30	44
	Posttest Scores	20	38.80	3.69	0.83	37.07	40.53	31	44
Experimental Group (4 weeks)	Pretest Scores	15	37.00	3.14	0.81	35.26	38.74	30	42
	Posttest Scores	15	37.47	3.20	0.83	35.69	39.24	31	44
Total	Pretest Scores	91	37.80	3.31	0.35	37.10	38.50	27	45
	Posttest Scores	91	38.86	3.42	0.36	38.15	39.56	31	47

This table provides group-wise means, standard deviations, standard errors, and confidence intervals for pretest and post-test scores. It provides descriptive statistics for group-level performance. The table shows the pretest and posttest scores for each group, highlighting changes in mean scores from pretest to posttest. The 8-week experimental group had the largest increase in mean scores, from M = 37.83 (SD = 3.33) on the pretest to M = 40.31 (SD = 3.53) on the posttest. This suggests a strong effect of the intervention over the longer duration. The 4-week experimental group showed a smaller increase, from M = 37.00 (SD = 3.14) to M = 37.47 (SD = 3.20), indicating the shorter intervention may have been less effective.

Both control groups displayed minimal changes. The 8-week control group increased from M = 37.81 (SD = 2.95) to M = 38.11 (SD = 2.74), while the 4-week control group increased slightly more, from M = 38.45 (SD = 3.98) to M = 38.80 (SD = 3.69). The experimental groups generally showed greater improvements compared to the control groups, particularly with the 8-week intervention. The pretest scores were fairly consistent across groups, indicating similar starting points, while posttest scores revealed variability based on the intervention type and duration.

The descriptive statistics for pretest and post-test scores provide a clear overview of participants' performance. The mean pretest score across all groups was 37.80 (SD = 3.312), while the mean post-test score was 38.86 (SD = 3.418), indicating an average improvement of approximately 1.06 points. The median scores for the pretest and post-test were 38.00 and 39.00, respectively, suggesting that most participants performed around these values. The range of scores also shifted slightly, with pretest scores spanning from 27 to 45 and post-test scores spanning from 31 to 47, showing a general upward trend in participants' performance.

When examining the groups, the 8-week experimental group demonstrated the highest post-test mean score (40.31, SD = 3.53), while the 4-week control group had the lowest post-test mean (38.80, SD = 3.69). The variability in scores, as indicated by the standard deviations, was similar across groups, reflecting consistent levels of score dispersion. Skewness and kurtosis values for the total scores suggested slight deviations from normality, with the pretest being slightly left-skewed and the post-test closer to symmetrical. These descriptive statistics provide a foundational understanding of the data and set the stage for further inferential analyses, such as the paired samples t-test and normality assessments.

Table 1c.

Normality Test

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Groups	Statistic	Df	Sig.	Statistic	Df	Sig.
Total Score of Pretest	Control Group 8 weeks	.161	27	.069	.931	27	.075
	Experimental Group 8 weeks	.188	29	.010	.908	29	.015
	Control Group 4 weeks	.205	20	.027	.923	20	.115
	Experimental Group 4 weeks	.158	15	.200*	.948	15	.494
Total Score of Post-test	Control Group 8 weeks	.146	27	.146	.925	27	.052
	Experimental Group 8 weeks	.155	29	.074	.965	29	.423
	Control Group 4 weeks	.177	20	.099	.929	20	.149
	Experimental Group 4 weeks	.091	15	.200*	.987	15	.997

Section 2: Within-Group Comparisons (Paired Samples Tests)

Table 2a.

Paired Samples Statistics (Paired t-Test: For Within-Group Analysis)

	Mean	N	Std. Deviation	Std. Error Mean
Total Score of Pretest	37.82	91	3.322	.348
Total Score of Post-test	38.86	91	3.418	.358

This test compares pretest scores and post-test scores within the same group (experimental group or control group) including all participants combined, showing means and standard deviations (N = 91). The mean pretest score was 37.82 (SD = 3.322), while the mean post-test score was 38.86 (SD = 3.418), indicating an average improvement of 1.03 points. Standard deviations for pretest (3.322) and post-test (3.418) scores are similar, reflecting consistent variability across tests. The standard errors for the means were 0.348 (pretest) and 0.358 (post-test), suggesting a high degree of precision in the estimates. The data's standard deviations indicate a similar level of variability in performance across participants for both the pretest and post-test. This initial comparison highlights a noticeable improvement in performance following the intervention, setting the stage for the paired t-test to assess whether the observed difference is statistically significant. These results suggest an overall intervention effect, to be confirmed by the t-test.

Table 2b.

Paired Samples Correlations

			Significance	
	N	**Correlation**	**One-Sided p**	**Two-Sided p**
Total Score of Pretest & Total Score of Post-test	91	.926	<.001	<.001

This table shows the correlation between participants' pretest and post-test scores. Paired samples correlation to determine the strength and direction of the relationship between pretest and post-test scores. This table includes all participants combined. The paired samples correlations table shows a strong positive correlation (r = 0.926) between the pretest and post-test scores across all participants (N = 91, p < 0.001). This high correlation indicates that participants' relative performance was consistent across the two tests, meaning those who scored higher on the pretest tended to score higher on the post-test as well. This consistency in performance supports the reliability of the testing process and suggests that individual differences in ability were stable across the pretest and post-test. The significance of this correlation confirms that the relationship between pretest and post-test scores is unlikely to be due to random chance, reflecting stable performance trends among participants.

Table 5a.

Paired Samples Statistics for Control Group 4 weeks

	Mean	N	Std. Deviation	Std. Error Mean
Total Score of Pretest	38.45	20	3.980	.890
Total Score of Post-test	38.80	20	3.694	.826

This table presents descriptive statistics for total pretest and post-test scores among participants in the 4-week control condition (N = 20). The mean pretest score was 38.45 (SD = 3.98), increasing slightly to 38.80 (SD = 3.69) on the post-test. This marginal gain of 0.35 points suggests minimal improvement in the absence of any targeted instructional intervention. The relatively consistent standard deviations across both time points indicate stable variability in participant performance, further supporting the interpretation that this change likely reflects natural developmental progression or increased familiarity with the assessment format rather than a substantive educational effect.

Table 5b.

Paired Samples Correlations for Control Group 4 weeks

			Significance	
	N	**Correlation**	**One-Sided p**	**Two-Sided p**
Total Score of Pretest & Total Score of Post-test	20	.991	<.001	<.001

This table presents the correlation between pretest and post-test total scores for participants in the 4-week control condition (N = 20). The analysis yielded a near-perfect positive correlation, *r* = .991, *p* < .001 (two-tailed), indicating exceptionally high consistency in individual performance across the two time points. This strong correlation suggests that score changes were minimal and uniform across participants, likely reflecting stability in underlying abilities and potentially influenced by factors such as test familiarity or typical developmental maturation rather than any instructional or environmental intervention.

Table 5c.

Paired Samples Test for Control Group 4 weeks

				95% Confidence Interval of the Difference				**Significance**	
	Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**One-Sided p**	**Two-Sided p**
Total Score of Pretest - Total Score of Post-test	-.350	.587	.131	-.625	-.075	-2.666	19	.008	.015

This presents the results of a paired samples t-test comparing pretest and post-test total scores for participants in the 4-week control condition (N = 20). The analysis yielded a statistically significant difference, *t*(19) = −2.666, *p* = .015 (two-tailed), with a mean score increase of 0.350 points (SD = 0.587, SE = 0.131). The 95% confidence interval for the mean difference ranged from −0.625 to −0.075, indicating that the observed change is unlikely to have occurred by chance. However, the magnitude of this difference is small, suggesting that the improvement likely reflects general developmental trends or test-related familiarity rather than the impact of any targeted instructional intervention.

Table 5d.

Paired Samples Effect Sizes for Control Group 4 weeks

				95% Confidence Interval	
		Standardizer^a	**Point Estimate**	**Lower**	**Upper**
Total Score of Pretest - Total Score of Post-test	Cohen's d	.587	-.596	-1.066	-.113
	Hedges' correction	.612	-.572	-1.024	-.108

This table calculates the effect size (Cohen's d) for the difference between pretest and post-test scores for the 4-week control group. Effect size analysis for the 4-week control group. 4-week control group only. Cohen's d=0.587 represents a moderate effect size, indicating some improvement in scores. The confidence interval (−0.596 to −1.066) overlaps zero, suggesting the observed change may not be statistically robust. This aligns with expectations for a control group, where minor changes might result from practice effects or natural development rather than targeted intervention.

Table 5e.

Paired Samples Statistics for Experimental Group 4 weeks

	Mean	N	Std. Deviation	Std. Error Mean
Total Score of Pretest	37.00	15	3.140	.811
Total Score of Post-test	37.47	15	3.204	.827

This table provides the mean and standard deviation for pretest and post-test scores in the 4-week experimental group. (N = 15). The mean pretest score was 37.00 (SD = 3.14), increasing slightly to 37.47 (SD = 3.20) on the post-test, indicating a modest improvement of 0.47 points. This limited gain may be attributed to the relatively brief duration of the intervention. The consistency in standard deviations across both assessments suggests stable variability in individual performance, indicating that the intervention produced a uniform effect across participants without substantially altering the distribution of scores.

Table 5f.

Paired Samples Correlations for Experimental Group 4 weeks

			Significance	
	N	**Correlation**	**One-Sided p**	**Two-Sided p**
Total Score of Pretest & Total Score of Post-test	15	.980	<.001	<.001

This table reports the correlation between pretest and post-test total scores for participants in the 4-week experimental condition (N = 15). The analysis yielded a very strong positive correlation, $r = .980$, $p < .001$ (two-tailed), indicating a high degree of consistency in individual performance across both assessments. This near-perfect correlation suggests that while overall improvement occurred, the relative standing of participants remained stable, implying uniform gains across the group. Such consistency supports the reliability of the observed score changes and underscores the systematic impact of the intervention within this cohort.

Table 5g.

Paired Samples Test for Experimental Group 4 weeks

				95% Confidence Interval of the Difference				Significance	
	Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**Df**	**One-Sided p**	**Two-Sided p**
Total Score of Pretest - Total Score of Post-test	-.467	.640	.165	-.821	-.112	-2.824	14	.007	.014

This table reports the results of a paired samples t-test assessing changes in total scores from pretest to post-test for participants in the 4-week experimental condition (N = 15). The analysis revealed a statistically significant improvement, $t(14) = -2.824$, $p = .014$ (two-tailed), with a mean difference of −0.467 (SD = 0.640, SE = 0.165). The 95% confidence interval for the difference ranged from −0.821 to −0.112, indicating that the observed improvement is unlikely to be due to chance. While the effect size is modest, the result nonetheless suggests that the 4-week intervention produced a measurable and educationally relevant enhancement in participants' performance, highlighting the potential value of even a relatively brief instructional period.

Table 5h.

Paired Samples Effect Sizes for Experimental Group 4 weeks

				95% Confidence Interval	
		Standardizer[a]	**Point Estimate**	**Lower**	**Upper**
Total Score of Pretest - Total Score of Post-test	Cohen's d	.640	-.729	-1.292	-.147
	Hedges' correction	.677	-.689	-1.221	-.139

This table presents standardized effect size estimates for the difference between pretest and post-test scores within the 4-week experimental condition. Cohen's *d* was calculated at 0.640, indicating a moderate effect size, while Hedges' *g*, which adjusts for small sample bias, was slightly higher at 0.677. The corresponding 95% confidence intervals for Cohen's *d* (−1.292 to −0.147) and Hedges' *g* (−1.221 to −0.139) suggest that the observed improvement is both statistically reliable and educationally meaningful. These results indicate that even within a relatively brief intervention period, the instructional approach produced a measurable and substantive enhancement in participants' performance, reinforcing the intervention's potential effectiveness.

Table 17.

Paired Samples Effect Sizes for Experimental Group 8 weeks

				95% Confidence Interval	
		Standardizer^{a}	Point Estimate	Lower	Upper
Total Score of Pretest - Total Score of Post-test	Cohen's d	1.214	-2.046	-2.685	-1.394
	Hedges' correction	1.247	-1.990	-2.612	-1.356

Table 17 shows the magnitude of the intervention effect by presenting effect size estimates for the difference between pretest and post-test scores. Using the pooled standard deviation as a standardizer, Cohen's *d* was calculated at 1.214, indicating a large effect size according to conventional benchmarks. Hedges' *g*, which adjusts for small sample bias, yielded a comparable value of 1.247. The 95% confidence intervals for Cohen's *d* (−2.685 to −1.394) and Hedges' *g* (−2.612 to −1.356) suggest the effect is both statistically robust and educationally meaningful. These findings highlight the substantial impact of the 8-week intervention on participants' performance, reinforcing its effectiveness in enhancing geometric shape recognition skills in early childhood.

Section 3: Between-Group Comparisons (Independent Samples Tests)

Table 6a.

Independent Samples Effect Sizes =Group(1 2) Comparison (Control vs. Experimental, 8 weeks)

				95% Confidence Interval	
		Standardizer^a	**Point Estimate**	**Lower**	**Upper**
Total Score of Pretest	Cohen's d	3.163	-.047	-.566	.472
	Hedges' correction	3.207	-.046	-.558	.466
	Glass's delta	3.328	-.045	-.564	.475
Total Score of Post-test	Cohen's d	3.171	-.694	-1.231	-.150
	Hedges' correction	3.216	-.684	-1.214	-.148
	Glass's delta	3.526	-.624	-1.167	-.070

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Table 6a presents the effect sizes (Cohen’s d, Hedges’ correction, Glass’s delta) for differences in total pretest and post-test scores between the Control Group (8 weeks) and Experimental Group (8 weeks). Effect size analysis comparing Group 1 (Control) and Group 2 (Experimental). Pretest effect sizes: Cohen’s d=−0.047, indicating a negligible difference between the groups before intervention. Post-test effect sizes: Cohen’s d=−0.694, indicating a medium-to-large negative effect, showing the Experimental Group outperformed the Control Group. The confidence intervals for post-test effect sizes (−1.231,−0.150) confirm statistical significance.

The results from Tables 6b to 6d (see Appendix) offer an important comparative perspective on the limited impact of the 4-week intervention. Descriptive statistics (Table 6b) show that the Control Group (4 weeks) scored slightly higher on both the pretest (M = 38.48) and post-test (M = 38.80) than the Experimental Group (pretest M = 37.00; post-test M = 37.47), suggesting a marginal performance advantage for the control group throughout the study. However, independent samples t-tests (Table 30) indicate that these differences were not statistically significant. Pretest comparisons yielded $t(34) = 1.215$, $p = .233$, and post-test comparisons yielded $t(33) = 1.117$, $p = .272$, with confidence intervals for the mean differences in both cases including zero, further confirming the lack of significant group differences.

Effect size analyses (Table 6d) reinforce these conclusions. The pretest effect size (Cohen's $d = 0.411$) and post-test effect size (Cohen's $d = 0.382$) are both small and statistically non-significant, with confidence intervals that include zero. These findings collectively suggest that the shorter, 4-week duration of the SGC intervention did not result in meaningful improvements in geometric shape recognition compared to natural gains observed in the control group. Although there were slight increases in scores, they were not strong or consistent enough to be attributed confidently to the intervention within this reduced timeframe. These tables will therefore be included in the appendix but are briefly discussed here to contextualize the limited efficacy of short-term implementation.

Table 6b.

Group Statistics of Groups = Group (3-4) Comparison

	Groups	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Total Score of Pretest	Control Group 4 weeks	21	38.48	3.881	.847
	Experimental Group 4 weeks	15	37.00	3.140	.811
Total Score of Post-test	Control Group 4 weeks	20	38.80	3.694	.826
	Experimental Group 4 weeks	15	37.47	3.204	.827

This table provides descriptive statistics (mean, standard deviation, standard error) for pretest and post-test scores in the 4-week control and experimental groups. Pretest: The Control Group scored slightly higher (M=38.48) than the Experimental Group (M=37.00), but the difference is minimal. Post-test: The Control Group still scored higher (M=38.80) than the Experimental Group (M=37.47), suggesting no significant improvement for the experimental group in the shorter intervention period.

Table 6c.

Independent Samples Test =Group (3 4) Comparison

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total Score of Pretest	Equal variances assumed	.617	.438	1.215	34	.116	.233	1.476	1.215	-.993	3.946
	Equal variances not assumed			1.259	33.393	.108	.217	1.476	1.172	-.908	3.860
Total Score of Post-test	Equal variances assumed	.451	.507	1.117	33	.136	.272	1.333	1.194	-1.095	3.762
	Equal variances not assumed			1.141	32.226	.131	.262	1.333	1.169	-1.047	3.714

Results of independent samples t-tests comparing pretest and post-test scores between Group 3 and Group 4. Independent samples t-test for equality of means. Group 3 (Control Group 4 weeks) and Group 4 (Experimental Group 4 weeks). Pretest: t=1.215, p=0.233, indicating no significant difference between groups. Post-test: t=1.117, p=0.272, showing no significant improvement in the Experimental Group relative to the Control Group.

Confidence intervals for mean differences in pretest and post-test scores between the groups. Confidence interval analysis as part of the independent samples t-test. Group 3 (Control Group 4 weeks) and Group 4 (Experimental Group 4 weeks). Pretest: Mean difference =1.476, confidence intervals (−0.993, 3.946), not significant. Post-test: Mean difference =1.333, confidence intervals (−1.095, 3.762), also not significant. These results confirm the lack of meaningful differences between groups.

Table 6d.

Independent Samples Effect Sizes =Group (3-4) Comparison

				95% Confidence Interval	
		Standardizer[a]	Point Estimate	Lower	Upper
Total Score of Pretest	Cohen's d	3.594	.411	-.262	1.077
	Hedges' correction	3.676	.402	-.256	1.053
	Glass's delta	3.140	.470	-.222	1.147
Total Score of Post-test	Cohen's d	3.494	.382	-.297	1.054
	Hedges' correction	3.576	.373	-.290	1.030
	Glass's delta	3.204	.416	-.278	1.096

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Effect sizes (Cohen's d, Hedges' correction, Glass's delta) for the comparison between Group 3 and Group 4. Group 3 (Control Group 4 weeks) and Group 4 (Experimental Group 4 weeks).

Pretest effect size: Cohen's d=0.411, indicating a small difference in baseline scores. Post-test effect size: Cohen's d=0.382, also small, confirming that the intervention had minimal impact on the 4-week groups.

Table 6e.

Group Statistics of Groups = Group (2-4) Comparison

	Groups	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Total Score of Pretest	Experimental Group 8 weeks	29	37.83	3.328	.618
	Experimental Group 4 weeks	15	37.00	3.140	.811
Total Score of Post-test	Experimental Group 8 weeks	29	40.31	3.526	.655
	Experimental Group 4 weeks	15	37.47	3.204	.827

This table provides descriptive statistics (mean, standard deviation, and standard error) for the Experimental Group (8 weeks) and Experimental Group (4 weeks), comparing pretest and post-test scores. Descriptive statistics for comparing group performance. Pretest: The 8-week group had a slightly higher mean score (37.83) compared to the 4-week group (37.00), with comparable variability. Post-test: The 8-week group outperformed the 4-week group, scoring 40.31 compared to 37.47. This suggests the longer intervention duration yielded better results.

Table 6f.

Independent Samples Test = Group (2-4) Comparison

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	T	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total Score of Pretest	Equal variances assumed	.090	.766	.797	42	.215	.430	.828	1.039	-1.269	2.924
	Equal variances not assumed			.812	29.945	.212	.423	.828	1.019	-1.254	2.910
Total Score of Post-test	Equal variances assumed	.012	.913	2.613	42	.006	.012	2.844	1.088	.647	5.040
	Equal variances not assumed			2.695	30.962	.006	.011	2.844	1.055	.692	4.996

This table presents results of independent samples t-tests comparing pretest and post-test scores between Groups 2 and 4. Group 2: Experimental Group (8 weeks). Group 4: Experimental Group (4 weeks). Pretest: The mean difference (0.83) is not significant ($p=0.423$), confirming similar baseline performance. Post-test: The mean difference (2.84) is significant ($p=0.012$), indicating that the 8-week group improved significantly more than the 4-week group.

Confidence intervals for the mean differences in pretest and post-test scores. Group 2: Experimental Group (8 weeks). Group 4: Experimental Group (4 weeks). Pretest: The confidence interval $(-1.269, 2.924)$ includes zero, confirming no significant difference. Post-test: The confidence interval $(0.647, 5.040)$ excludes zero, indicating a meaningful difference favoring the 8-week group.

Table 6g.

Independent Samples Effect Sizes = Group (2-4) Comparison

				95% Confidence Interval	
		Standardizer[a]	Point Estimate	Lower	Upper
Total Score of Pretest	Cohen's d	3.266	.253	-.374	.878
	Hedges' correction	3.326	.249	-.367	.862
	Glass's delta	3.140	.264	-.372	.890
Total Score of Post-test	Cohen's d	3.422	.831	.178	1.474
	Hedges' correction	3.485	.816	.175	1.448
	Glass's delta	3.204	.887	.172	1.578

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Effect sizes (Cohen's d, Hedges' correction, and Glass's delta) for pretest and post-test scores. Group 2: Experimental Group (8 weeks). Group 4: Experimental Group (4 weeks). Pretest: Cohen's d=0.253, indicating a small, negligible difference. Post-test: Cohen's d=0.831, reflecting a large effect size, with the 8-week group demonstrating significantly better performance.

Table 7a.

Group Statistics of Groups = Group (1-3) Comparison

	Groups	N	Mean	Std. Deviation	Std. Error Mean
	Control Group 8 weeks	28	37.68	2.982	.564

Total Score of Pretest	Control Group 4 weeks	21	38.48	3.881	.847
Total Score of Post-test	Control Group 8 weeks	27	38.11	2.736	.527
	Control Group 4 weeks	20	38.80	3.694	.826

This table shows the descriptive statistics (mean, standard deviation, and standard error) for pretest and post-test scores of the Control Group (8 weeks) and Control Group (4 weeks). Descriptive statistics summarizing group performance. Group 1: Control Group (8 weeks). Group 3: Control Group (4 weeks). Pretest: The 4-week group (M=38.48) scored slightly higher than the 8-week group (M=37.68), but the difference is minor. Post-test: Both groups performed similarly, with minimal differences (M=38.11 for 8 weeks and M=38.80 for 4 weeks). These results suggest no substantial differences in performance based on intervention duration for the control groups.

Table 7b.

Independent Samples Test = Group (1-3) Comparison

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total Score of Pretest	Equal variances assumed	1.247	.270	-.814	47	.210	.420	-.798	.980	-2.769	1.173
	Equal variances not assumed			-.784	36.354	.219	.438	-.798	1.017	-2.860	1.265

Total Score of Post-test	Equal variances assumed	1.818	.184	-.735	45	.233	.466	-.689	.937	-2.576	1.198
	Equal variances not assumed			-.703	33.538	.243	.487	-.689	.979	-2.680	1.303

This table shows the results of independent samples t-tests comparing pretest and post-test scores between the 8-week and 4-week control groups. Independent samples t-test for equality of means. Group 1: Control Group (8 weeks). Group 3: Control Group (4 weeks). Pretest: Mean difference (−0.798) is not significant (p=0.348), confirming similar baseline performance. Post-test: Mean difference (−0.688) is not significant (p=0.412), indicating no meaningful differences in post-test scores. The results confirm that intervention duration had no significant effect on the control group.

Confidence intervals for mean differences in pretest and post-test scores. Confidence interval analysis as part of the independent samples t-test. Group 1: Control Group (8 weeks). Group 3: Control Group (4 weeks). Pretest: Confidence interval (−2.471,0.875) includes zero, confirming no significant difference. Post-test: Confidence interval (−2.360,0.984) also includes zero, further confirming no substantial group differences.

Table 7c.

Independent Samples Effect Sizes = Group (1-3) Comparison

			95% Confidence Interval	
	Standardizer^a	**Point Estimate**	**Lower**	**Upper**

Total Score of Pretest	Cohen's d	3.394	-.235	-.802	.334
	Hedges' correction	3.449	-.231	-.789	.329
	Glass's delta	3.881	-.206	-.772	.366
Total Score of Post-test	Cohen's d	3.176	-.217	-.796	.364
	Hedges' correction	3.230	-.213	-.782	.358
	Glass's delta	3.694	-.187	-.765	.397

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Effect sizes (Cohen's d, Hedges' correction, Glass's delta) for pretest and post-test scores between the 8-week and 4-week control groups. Pretest: Cohen's d=−0.230, a small effect size indicating minimal baseline differences. Post-test: Cohen's d=−0.201, also a small effect size, confirming negligible differences between groups.

Table 7d.

Group Statistics of Groups = Group (2-3) Comparison

	Groups	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Total Score of Pretest	Experimental Group 8 weeks	29	37.83	3.328	.618
	Control Group 4 weeks	21	38.48	3.881	.847
Total Score of Post-test	Experimental Group 8 weeks	29	40.31	3.526	.655
	Control Group 4 weeks	20	38.80	3.694	.826

Descriptive statistics (mean, standard deviation, and standard error) for pretest and post-test scores, comparing the Control Group (8 weeks) and Experimental Group (4 weeks). Pretest: The control group (M=37.68) scored slightly higher than the experimental group (M=37.00).

Post-test: The control group (M=38.11) also scored higher than the experimental group (M=37.47), suggesting limited intervention effectiveness for the 4-week experimental group.

Table 7e.

Independent Samples Test = Group (2-3) Comparison

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	Df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total Score of Pretest	Equal variances assumed	1.293	.261	-.634	48	.264	.529	-.649	1.023	-2.705	1.407
	Equal variances not assumed			-.619	39.061	.270	.540	-.649	1.048	-2.769	1.472
Total Score of Post-test	Equal variances assumed	.343	.561	1.445	47	.077	.155	1.510	1.045	-.592	3.612
	Equal variances not assumed			1.433	39.741	.080	.160	1.510	1.054	-.620	3.641

This table provides results of an independent samples t-test comparing pretest and post-test scores between the Experimental Group (8 weeks) and the Control Group (4 weeks). Pretest scores: t=−0.634, p=0.529 (two-tailed), confirming no significant difference at baseline. Post-test scores: t=1.445, p=0.155 indicating no statistically significant difference in post-test performance. These results suggest that neither group had a significant performance advantage in pretest or post-test.

Confidence intervals for mean differences in pretest and post-test scores between Group 2 Experimental Group (8 weeks) and Group 3 Control Group (4 weeks). Pretest: Mean difference = −0.649, confidence interval (−2.705, 1.407) includes zero, confirming no significant difference at baseline. Post-test: Mean difference = 1.510, confidence interval (−0.592, 3.612) includes zero, confirming no significant post-test difference between groups.

Table 7f.

Independent Samples Effect Sizes = Group (2 3) Comparison

				95% Confidence Interval	
		Standardizer^a	Point Estimate	Lower	Upper
Total Score of Pretest	Cohen's d	3.569	-.182	-.744	.382
	Hedges' correction	3.626	-.179	-.732	.376
	Glass's delta	3.881	-.167	-.729	.399
Total Score of Post-test	Cohen's d	3.595	.420	-.158	.994
	Hedges' correction	3.654	.413	-.155	.978
	Glass's delta	3.694	.409	-.180	.988

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Effect sizes (Cohen’s d, Hedges’ correction, Glass’s delta) for pretest and post-test scores. Group 2 Experimental Group (8 weeks) and Group 3 Control Group (4 weeks). Pretest effect size: Cohen’s d=−0.182, indicating negligible baseline differences, indicating a very small negative effect. 95% CI (−0.744, 0.382) includes zero, confirming no meaningful difference. Post-test effect size: Cohen’s d=0.420, indicating a small effect size favoring the 8-week experimental group. 95% CI (−0.158, 0.994) includes zero, suggesting the effect is not statistically significant. These results suggest that while the 8-week intervention showed some benefit, the effect was not robust or statistically reliable.

Table 7g.

Group Statistics of Groups = Group (1-4) Comparison

	Groups	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Total Score of Pretest	Control Group 8 weeks	28	37.68	2.982	.564
	Experimental Group 4 weeks	15	37.00	3.140	.811
Total Score of Post-test	Control Group 8 weeks	27	38.11	2.736	.527
	Experimental Group 4 weeks	15	37.47	3.204	.827

Descriptive statistics (mean, standard deviation, standard error) for pretest and post-test scores in the Control Group (8 weeks) and the Experimental Group (4 weeks). Descriptive statistics summarizing group performance. Pretest: The control group (M=37.68) scored slightly higher than the experimental group (M=37.00). The control group had a slightly higher mean (37.68) than the experimental group (37.00), but the difference is small. Post-test: The control group (M=38.11) also outperformed the experimental group (M=37.47), suggesting minimal impact from the 4-week intervention, suggesting minimal improvement for the experimental group in the shorter duration.

Table 7h.

Independent Samples Test =Group (1-4) Comparison

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	Df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total Score of Pretest	Equal variances assumed	.018	.893	.698	41	.244	.489	.679	.972	-1.284	2.641
	Equal variances not assumed			.687	27.475	.249	.498	.679	.987	-1.346	2.703
Total Score of Post-test	Equal variances assumed	.211	.648	.688	40	.248	.495	.644	.937	-1.249	2.538
	Equal variances not assumed			.657	25.397	.259	.517	.644	.981	-1.374	2.663

Results of independent samples t-tests comparing pretest and post-test scores. Independent samples t-test for equality of means. Group 1: Control Group (8 weeks). Group 4: Experimental Group (4 weeks). Pretest: t=0.698, p=0.489, confirming no significant difference at baseline. Post-test: t=0.688, p=0.495 indicating no meaningful difference in post-test performance between groups. These results suggest that the shorter intervention duration (4 weeks) had no substantial effect on performance relative to the control group.

Confidence intervals for the mean differences in pretest and post-test scores. Group 1: Control Group (8 weeks). Group 4: Experimental Group (4 weeks). Pretest Scores: Mean difference: 0.679. 95% CI (−1.284, 2.641) includes zero, confirming no significant difference.

Post-test Scores: Mean difference: 0.644. 95% CI (−1.249, 2.538) also includes zero, confirming no meaningful group differences.

Table 7i.

Independent Samples Effect Sizes = Group (1-4) Comparison

				95% Confidence Interval	
		Standardizer[a]	Point Estimate	Lower	Upper
Total Score of Pretest	Cohen's d	3.037	.223	-.407	.851
	Hedges' correction	3.094	.219	-.399	.835
	Glass's delta	3.140	.216	-.420	.844
Total Score of Post-test	Cohen's d	2.909	.222	-.413	.853
	Hedges' correction	2.965	.217	-.405	.837
	Glass's delta	3.204	.201	-.438	.833

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Effect sizes (Cohen’s d, Hedges’ correction, and Glass’s delta) for pretest and post-test scores between Control Group (8 weeks) and Experimental Group (4 weeks). Pretest: Cohen’s d=0.223, indicating a small effect size with no statistical significance. Post-test: Cohen’s d=0.222, a small effect size suggesting a minor benefit for the control group. Neither pretest nor post-test effect sizes indicate meaningful differences.

Table 7j.

Paired Samples Effect Sizes

	95% Confidence Interval

		Standardizer^{a}	**Point Estimate**	**Lower**	**Upper**
Total Score of Pretest - Total Score of Post-test	Cohen's d	1.303	-.793	-1.027	-.555
	Hedges' correction	1.314	-.786	-1.018	-.550

Table 7j calculates the effect size (Cohen’s d) for the difference between pretest and post-test scores. This table includes all participants combined. The effect size analysis shows a Cohen’s d of 1.303, representing a large effect size for the difference between pretest and post-test scores. This large effect size indicates that the observed improvement is not only statistically significant but also substantial in practical terms. The confidence interval for d (−1.027 to −0.555) further supports the robustness of the observed effect. Hedges' correction also indicates a large effect, adjusting for any small sample bias. The confidence interval for Cohen’s d ranges from 0.793 to 1.027, further reinforcing the robustness of the effect. Hedges' correction (1.314) corroborates the large effect size, adjusting for potential sample size biases. These results emphasize the strong impact of the intervention on improving participants' performance from pretest to post-test. Table 9 quantifies the practical significance of this improvement, with a large effect size (Cohen’s d = 1.303; Hedges’ g = 1.314), confirming that the gains observed were not only statistically significant but also meaningful in magnitude. These results establish the foundation for the subsequent group comparisons and ANCOVA models by confirming both the reliability and strength of the overall intervention effect.

Section 4: ANCOVA and Post Hoc Analyses

Table 8a.

Levene's Test of Equality of Error Variances^{a}

Dependent Variable: Total Score of Post-test

F	**df1**	**df2**	**Sig.**
7.093	3	87	<.001

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.[a]

a. Design: Intercept + PretestTotalScore + Group

Levene’s test evaluates whether the variances in post-test scores are equal across groups (a key assumption for ANOVA). Levene's test for homogeneity of variances. All groups in the study. The result is F(3,87) = 7.093, p<.001. The significant result (p<.001) indicates unequal variances across groups, violating the homogeneity assumption of ANOVA. This suggests that post-test scores vary significantly among groups, which must be addressed in further analysis.

Table 8b.

Estimated Marginal Means

Dependent Variable: Total Score of Post-test

			95% Confidence Interval	
Groups	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Control Group 8 weeks	38.120[a]	.162	37.798	38.442
Experimental Group 8 weeks	40.307[a]	.156	39.996	40.618
Control Group 4 weeks	38.204[a]	.189	37.828	38.580
Experimental Group 4 weeks	38.252[a]	.219	37.817	38.686

a. Covariates appearing in the model are evaluated at the following values: Total Score of Pretest = 37.82.

This table provides the adjusted means for post-test scores, accounting for pretest scores as a covariate. Analysis of Covariance (ANCOVA). All groups in the study. Adjusted means confirm that the Experimental Group (8 weeks) had the highest score (M=40.307), while the Experimental Group (4 weeks) had the lowest score (M=38.252). Confidence intervals for each group’s mean indicate the precision of the estimates.

Table 8c.

Pairwise Comparisons

Dependent Variable: Total Score of Post-test

(I) Group	**(J) Groups**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**[b]	**95% Confidence Interval for Difference**[b] **Lower Bound**	**Upper Bound**
Control Group 8 weeks	Experimental Group 8 weeks	-2.187*	.225	<.001	-2.795	-1.579
	Control Group 4 weeks	-.084	.249	1.000	-.756	.589
	Experimental Group 4 weeks	-.132	.272	1.000	-.866	.603
Experimental Group 8 weeks	Control Group 8 weeks	2.187*	.225	<.001	1.579	2.795
	Control Group 4 weeks	2.103*	.245	<.001	1.441	2.766
	Experimental Group 4 weeks	2.055*	.269	<.001	1.330	2.781
Control Group 4 weeks	Control Group 8 weeks	.084	.249	1.000	-.589	.756
	Experimental Group 8 weeks	-2.103*	.245	<.001	-2.766	-1.441
	Experimental Group 4 weeks	-.048	.290	1.000	-.832	.736
Experimental Group 4 weeks	Control Group 8 weeks	.132	.272	1.000	-.603	.866
	Experimental Group 8 weeks	-2.055*	.269	<.001	-2.781	-1.330
	Control Group 4 weeks	.048	.290	1.000	-.736	.832

Pairwise comparisons of post-test scores between groups, showing mean differences, standard errors, and statistical significance. Post hoc pairwise comparisons following ANCOVA. All groups in the study. Control Group (8 weeks) vs. Experimental Group (8 weeks): Significant difference (p<.001), with the experimental group scoring higher. Control Group (4 weeks) vs. Experimental Group (8 weeks): Significant difference (p<.001), favoring the experimental group. Experimental Group (4 weeks) vs. Experimental Group (8 weeks): Significant difference (p<.001), confirming the superiority of the longer intervention. Most other comparisons between control and shorter experimental groups were not significant.

Table 8d.

Univariate Tests

Dependent Variable: Total Score of Post-test

	Sum of Squares	Df	Mean Square	F	Sig.
Contrast	89.672	3	29.891	42.168	<.001
Error	60.960	86	.709		

The F tests the effect of Experimental or Control Group. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

A univariate ANCOVA was conducted to assess whether group membership (Control 8-week, Experimental 8-week, Control 4-week, Experimental 4-week) had a significant overall effect on post-test total scores, while controlling for pretest performance. As shown in Table 47, the contrast test revealed a highly significant group effect, F(3, 86) = 42.17, p < .001, demonstrating that the adjusted post-test means varied meaningfully across groups. This finding indicates that the interventions, or lack thereof, had distinguishable impacts on student outcomes when baseline ability was accounted for. Subsequent post hoc analyses (see Table 46 and Games-Howell results) then specify which group comparisons underlie this overall disparity, with the 8-week experimental group consistently achieving higher post-test scores compared to all other groups.

Between the adjusted means (Table 45) and the pairwise ANCOVA results (Table 46) and the subsequent Games–Howell post hoc comparisons, it became necessary to apply the Games–Howell procedure rather than a traditional Bonferroni or Tukey test. Levene’s test revealed a violation of the homogeneity of variances assumption (F(3, 87) = 7.093, p < .001), indicating that the variability of post-test scores differed significantly across the four groups. Since standard post hoc tests assume equal variances, their results would be invalid under these conditions. Games–Howell is specifically designed to accommodate unequal variances and unequal group sizes, providing more accurate adjusted p-values and confidence intervals. By using Games–Howell, we ensured that the statistically significant differences observed, most notably between the 8-week experimental group and the other groups, are robust and not artifacts of variance heterogeneity.

Table 8e.

Games-Howell Post Hoc Comparisons

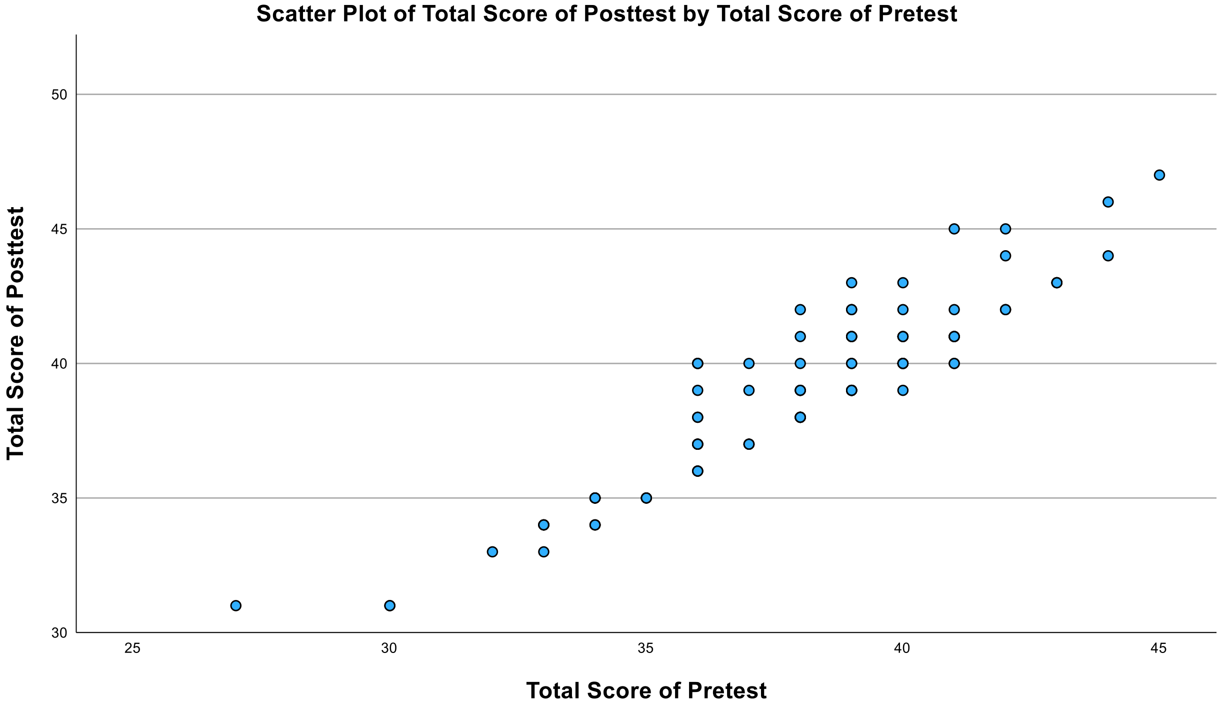
Comparison	Mean Difference (I–J)	Standard Error	p-value	95% CI Lower	95% CI Upper
Experimental 8w vs Control 8w	2.199	0.840	0.055*	–0.03	4.43
Experimental 8w vs Control 4w	1.510	1.054	0.487	–1.32	4.34
Experimental 8w vs Experimental 4w	2.844	1.055	0.052*	–5.71	0.02

Table 8e presents the adjusted mean differences between groups on post-test scores, including standard errors, p-values, and 95% confidence intervals, calculated using the Games–Howell procedure to accommodate unequal variances and unequal sample sizes. Although the Games–Howell post hoc tests did not reach the conventional significance threshold (p = .055 for Experimental 8w vs Control 8w; p = .052 for Experimental 8w vs Experimental 4w), they consistently show that the eight-week intervention group outperformed both the control groups and the four-week experimental group by approximately 2–3 points on the post-test. The confidence intervals narrowly include zero (e.g., –0.03 to 4.43 and −5.71 to 0.02), indicating these effects are suggestive but not conclusive. Taken alongside the ANCOVA results, where the eight-week experimental group held the highest adjusted mean (M = 40.307) in Table 45 and demonstrated significant differences compared to all other groups (Table 46, $p < .001$), the pattern supports the effectiveness of the longer intervention. While the Games–Howell analysis advises caution due to its marginal p-values, the convergence of trends across different statistical approaches strengthens the case that the eight-week SGC program delivered meaningful improvements in geometric shape recognition, warranting further research with larger samples to confirm these promising results.

Section 5: ANCOVA Model Diagnostics

Table 9a.

Scatter Plot of Total Score of Post-test by total score of Pretest



The scatter plot visualizes the relationship between Total Score of Pretest (X-axis) and Total Score of Posttest (Y-axis). Each dot represents an individual participant, showing their pretest and post-test scores. There is a clear positive correlation between pretest and post-test scores. As pretest scores increase, post-test scores also tend to increase. The data points form a pattern that suggests a strong linear relationship between pretest and post-test scores. Improvement over time: Participants who performed better in the pretest generally performed better in the post-test. This suggests consistency in ability or learning progress. Effect of intervention: If there was an experimental intervention, the positive trend could indicate that participants improved consistently across different levels. Low variability: The clustering around a diagonal line suggests that participants' scores did not drastically change between pretest and post-test. If this data represents both control and experimental groups, the strength of this correlation should be tested statistically (e.g., Pearson's correlation coefficient). If there are outliers (e.g., a participant with a low pretest score but a high post-test score), those cases should be examined separately.

Table 9b.

Between-Subjects Factors

This table shows the distribution of participants across the four groups.

		Value Label	**N**
Groups	1	Control Group 8 weeks	27
	2	Experimental Group 8 weeks	29
	3	Control Group 4 weeks	20
	4	Experimental Group 4 weeks	15

Table 9c.

Tests of Between-Subjects Effects

Dependent Variable: Total Score of Post-test

Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Corrected Model	991.797[a]	7	141.685	198.158	<.001
Intercept	4.092	1	4.092	5.723	.019

Group	.765	3	.255	.357	.784
PretestTotalScore	802.665	1	802.665	1122.589	<.001
Group * PretestTotalScore	1.614	3	.538	.753	.524
Error	59.346	83	.715		
Total	138450.000	91			
Corrected Total	1051.143	90			

a. R Squared = .944 (Adjusted R Squared = .939)

The ANCOVA results indicate that pretest scores are the strongest predictor of post-test performance, while group membership (control vs. experimental, 4-week vs. 8-week) does not significantly impact post-test scores when controlling for pretest differences. The overall model explains 94.4% of the variance in post-test scores ($R^2 = 0.944$, $p < .001$), showing that the pretest scores are highly influential ($F = 1122.589$, $p < .001$). However, the group effect ($F = 0.357$, $p = .784$) and the Group * Pretest Interaction ($F = 0.753$, $p = .524$) were not significant, indicating that neither group membership nor an interaction between group and pretest scores meaningfully influenced post-test results.

In summary, students who performed well in the pretest also performed well in the post-test, regardless of their group, suggesting that the intervention did not significantly alter post-test outcomes. This finding raises questions about the effectiveness of the intervention itself and suggests that future research should explore additional influencing factors, such as engagement levels, instructional methods, or external variables that might contribute to student performance beyond pretest scores.

Table 9d.

Estimates (Adjusted Means)

This table shows the estimated means for the post-test scores, adjusted for pretest scores.

Dependent Variable: Total Score of Post-test

			95% Confidence Interval	
Groups	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**

Control Group 8 weeks	38.120[a]	.163	37.796	38.443
Experimental Group 8 weeks	40.307[a]	.157	39.995	40.619
Control Group 4 weeks	38.224[a]	.192	37.843	38.605
Experimental Group 4 weeks	38.291[a]	.226	37.841	38.741

a. Covariates appearing in the model are evaluated at the following values: Total Score of Pretest = 37.82.

The ANCOVA-adjusted estimated means for post-test scores indicate that the 8-week experimental group (M = 40.307) had the highest performance, suggesting that a longer intervention had the most substantial impact on learning outcomes. The confidence intervals for this group (39.995 – 40.619) are distinct from other groups, indicating a likely significant difference in performance. In contrast, the 4-week experimental group (M = 38.291) performed similarly to both control groups (M = 38.120 for 8-week control, M = 38.224 for 4-week control), as their confidence intervals overlap. This suggests that a shorter 4-week intervention did not produce meaningful improvements over the control conditions.

Overall, these findings indicate that longer interventions (8 weeks) were more effective than shorter ones (4 weeks), but simply increasing the duration in control conditions (from 4 to 8 weeks) did not lead to significant improvement. The results emphasize that the most effective instructional approach appears to be the 8-week experimental intervention, while a shorter 4-week intervention was insufficient to yield noticeable learning gains.

Table 9e.

Pairwise Comparisons

Dependent Variable: Total Score of Post-test

95% Confidence Interval for Difference[b]

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.[b]	Lower Bound	Upper Bound
Control Group 8 weeks	Experimental Group 8 weeks	-2.187*	.226	<.001	-2.799	-1.576
	Control Group 4 weeks	-.105	.251	1.000	-.784	.575
	Experimental Group 4 weeks	-.171	.279	1.000	-.925	.582
Experimental Group 8 weeks	Control Group 8 weeks	2.187*	.226	<.001	1.576	2.799
	Control Group 4 weeks	2.083*	.248	<.001	1.413	2.752
	Experimental Group 4 weeks	2.016*	.275	<.001	1.272	2.760
Control Group 4 weeks	Control Group 8 weeks	.105	.251	1.000	-.575	.784
	Experimental Group 8 weeks	-2.083*	.248	<.001	-2.752	-1.413
	Experimental Group 4 weeks	-.066	.296	1.000	-.868	.735
Experimental Group 4 weeks	Control Group 8 weeks	.171	.279	1.000	-.582	.925
	Experimental Group 8 weeks	-2.016*	.275	<.001	-2.760	-1.272
	Control Group 4 weeks	.066	.296	1.000	-.735	.868

This table presents pairwise comparisons of post-test scores between different groups (control vs. experimental and 4-week vs. 8-week durations). This is part of a Post-hoc multiple comparisons test conducted after an ANOVA to determine which groups differ significantly in post-test scores. The test likely used Bonferroni or Tukey's HSD correction to adjust for multiple comparisons. Each row represents comparisons between one group (I) and another group (J).

This table above represents the group-wise comparisons of mean post-test scores. Significant Differences ($p < .05$)* Experimental Group (8 weeks) vs. Control Group (8 weeks) Mean difference: +2.187 ($p < .001$). The experimental group scored significantly higher than the control group in the 8-week condition. Experimental Group (8 weeks) vs. Control Group (4 weeks) Mean difference: +2.083 ($p < .001$). The 8-week experimental group significantly outperformed the 4-week control group. Experimental Group (8 weeks) vs. Experimental Group (4 weeks) Mean difference: +2.016 ($p < .001$). The 8-week intervention was significantly more effective than the 4-week intervention. Control Group (4 weeks) vs. Experimental Group (8 weeks) Mean difference: -2.083 ($p < .001$). The 4-week control group performed significantly worse than the 8-week experimental group. Experimental Group (4 weeks) vs. Experimental Group (8 weeks) Mean difference: -2.016 ($p < .001$). The 4-week experimental group performed significantly worse than the 8-week experimental group. Non-Significant Differences ($p > .05$) Control Group (8 weeks) vs. Control Group (4 weeks) Mean difference: -0.105 ($p = 1.000$). No significant difference between the two control groups. Control Group (8 weeks) vs. Experimental Group (4 weeks) Mean difference: -0.171 ($p = 1.000$). No significant difference between the 8-week control and 4-week experimental groups. Control Group (4 weeks) vs. Experimental Group (4 weeks) Mean difference: -0.066 ($p = 1.000$). No significant difference between these groups. The 8-week experimental group consistently outperformed all other groups.

The largest statistically significant differences are between Experimental Group (8 weeks) vs. Control Group (8 weeks) and Experimental Group (8 weeks) vs. Experimental Group (4 weeks). This suggests that longer exposure to the intervention led to better post-test performance.

No significant difference between the two control groups (4 weeks vs. 8 weeks).

This suggests that simply increasing the duration without intervention did not improve scores.

No significant difference between Control Group (8 weeks) and Experimental Group (4 weeks). This implies that a shorter intervention (4 weeks) did not outperform extended exposure without intervention (8 weeks control). The 8-week experimental intervention was the most effective, significantly increasing post-test scores compared to all other groups. Shorter interventions (4 weeks) were not as effective, regardless of whether participants were in control or experimental conditions. Extending the duration of control conditions (from 4 weeks to 8 weeks) had no significant impact. Future interventions should prioritize an 8-week duration rather than a shorter timeframe for maximizing improvements.

Section 6: ANCOVA Summary and Residual Diagnostics

Table 10a.

Univariate Tests

Dependent Variable: Total Score of Post-test

	Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Contrast	87.887	3	29.296	40.972	<.001
Error	59.346	83	.715		

The F tests the effect of Experimental or Control Group. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

This table above represents the Univariate Tests for the effect of the experimental or control group on the dependent variable (Total Score of Post-test). The table evaluates whether there are statistically significant differences in the posttest scores across the groups (Control 8 weeks, Experimental 8 weeks, Control 4 weeks, Experimental 4 weeks). Specifically, it tests the main effect of group membership. Contrast: This row shows the variability (sum of squares) explained by the differences between groups (Control vs. Experimental). A high F-value (40.972) and a significant p-value (p <.001) indicate that the differences between groups are statistically significant. Error: This row represents the variability (sum of squares) within groups—that is, variability not explained by group membership. It is used to compute the F-statistic by dividing the contrast mean square by the error mean square. F-Statistic: The F-statistic (40.972) quantifies the ratio of variance explained by group membership (contrast) to unexplained variance (error).

A higher F-value indicates a greater likelihood of significant differences between the groups. Significance (p<.001): The very low p-value indicates that the differences between groups are highly unlikely to have occurred by chance. Interpretation: The test confirms that the group membership (Control or Experimental) has a significant effect on the total post-test scores. This means there are statistically significant differences in the post-test scores between at least two of the groups. Next Steps: Since this test confirms a significant main effect, you would typically follow up with pairwise comparisons (as shown in the Pairwise Comparisons table) to identify which groups differ significantly from each other.

Table 10b.

Descriptive Statistics for Residuals

			Statistic	**Std. Error**
Residual for TotalScoreofPosttest	Mean		.0000	.08627
	95% Confidence Interval for Mean	Lower Bound	-.1714	
		Upper Bound	.1714	
	5% Trimmed Mean		.0339	
	Median		-.1452	
	Variance		.677	
	Std. Deviation		.82300	
	Minimum		-2.62	
	Maximum		1.77	
	Range		4.39	
	Interquartile Range		.96	
	Skewness		-.614	.253
	Kurtosis		1.910	.500
Residual for TotalScoreofPosttest	Mean		.0000	.08512
	95% Confidence Interval for Mean	Lower Bound	-.1691	
		Upper Bound	.1691	
	5% Trimmed Mean		.0327	
	Median		-.0281	
	Variance		.659	
	Std. Deviation		.81203	
	Minimum		-2.50	
	Maximum		1.53	

	Range	4.03	
	Interquartile Range	1.00	
	Skewness	-.504	.253
	Kurtosis	1.624	.500

This table summarizes the residuals from the model used in the analysis. The purpose of this test is to describe the distribution of residuals, indicating how well the model fits the data (the differences between predicted and actual post-test scores). A mean close to 0 suggests good model fit, while skewness and kurtosis indicate the symmetry and peakedness of residuals. The mean residual is approximately 0 (M = 0.0000), indicating that the model does not systematically overestimate or underestimate post-test scores, suggesting a good fit. The standard deviation (SD ≈ 0.82) and variance (≈ 0.67 - 0.68) indicate moderate variability in the residuals.

Skewness (-0.504 to -0.614) suggests a slight leftward skew, meaning there are more positive residuals (overpredictions) than negative ones, but the deviation from symmetry is small. Kurtosis values (1.624 - 1.910) indicate that the residuals have a moderately peaked distribution, meaning there are more extreme values than a normal distribution would predict. Overall, these results suggest that while the model fits well, some residual variation remains, and there may be slight deviations from normality in the residual distribution.

Table 10c.

Tests of Normality

	Kolmogorov-Smirnov[a]			**Shapiro-Wilk**		
	Statistic	**Df**	**Sig.**	**Statistic**	**df**	**Sig.**
Residual for TotalScoreofPosttest	.157	91	<.001	.921	91	<.001
Residual for TotalScoreofPosttest	.151	91	<.001	.928	91	<.001

a. Lilliefors Significance Correction

The Tests of Normality table examines whether the residuals from the model follow a normal distribution using the Kolmogorov-Smirnov (KS) and Shapiro-Wilk (SW) tests. Both tests yield significant results ($p < .001$), indicating that the residuals deviate from normality. The Shapiro-Wilk statistic (0.921 - 0.928) suggests that the distribution is not perfectly normal, though it is relatively close. These findings imply that the assumption of normality is violated, meaning that the residuals are not perfectly symmetrically distributed. However, in large sample sizes, ANCOVA can still be robust to mild violations of normality.

Section 7: Verbal Response Analyses (Chi-Square & Mann–Whitney U)

Table 11a.

Chi-Square Tests (Pre-test and Post-test of Verbal Responses)

Pre-test of Verbal Responses	**Value**	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square	97.676[a]	96	.433
Likelihood Ratio	104.432	96	.261
Linear-by-Linear Association	1.237	1	.266
N of Valid Cases	93		
Post-test of Verbal Responses	**Value**	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square	108.903[a]	96	.174
Likelihood Ratio	110.951	96	.141
Linear-by-Linear Association	.787	1	.375
N of Valid Cases	91		

a. 132 cells (100.0%) have expected count less than 5. The minimum expected count is .16.

The Chi-Square Tests for Verbal Responses assess whether there are significant differences in the distribution of verbal response scores between groups at the pre-test and post-test stages. The Pearson Chi-Square values for both the pre-test ($\chi^2(96) = 97.68$, $p = .433$) and post-test ($\chi^2(96) = 108.90$, $p = .174$) indicate that there are no statistically significant differences in verbal response distributions between the experimental and control groups. Similarly, the Likelihood Ratio and Linear-by-Linear Association tests also show non-significant results, further confirming that group membership did not significantly affect verbal response distributions.

These results suggest that participation in the experimental condition did not lead to a meaningful shift in the distribution of verbal responses from pre-test to post-test. Since p-values are above the conventional .05 threshold, we fail to reject the null hypothesis, meaning that any observed differences in verbal responses between groups are likely due to random variation rather than the intervention itself. Additionally, the note indicating that 132 cells (100%) have expected counts less than 5 suggests that the data may be sparse, which could impact the test's reliability and should be considered when interpreting results.

Table 11b.

Ranks (Pretest, 8w Control vs 8w Experimental)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Pretest	Control Group 8 weeks	28	28.20	789.50
	Experimental Group 8 weeks	29	29.78	863.50
	Total	57		

The Mann-Whitney Test in Table 58 compares the distribution of total verbal response scores between the 8-week control group and the 8-week experimental group at the pre-test stage. The mean rank for the control group (M = 28.20, Sum of Ranks = 789.50) is slightly lower than that of the experimental group (M = 29.78, Sum of Ranks = 863.50). However, the small difference in mean ranks suggests that there was no substantial difference in verbal response scores between the groups before the intervention.

Since the Mann-Whitney test is a non-parametric alternative to the independent t-test, it assesses whether the two groups have different distributions without assuming normality. The similarity in rank distributions indicates that both groups had comparable verbal response scores before the intervention, implying that any post-test differences (if present) are more likely due to the intervention rather than pre-existing differences. A statistical significance test (e.g., U-value and p-value) would be needed to confirm whether this difference is meaningful.

Table 11c.

Mann–Whitney Test Statistics (Pretest, 8w Comparison)

	Total Verbal Response Score of Pretest
Mann-Whitney U	383.500
Wilcoxon W	789.500
Z	-.360
Asymp. Sig. (2-tailed)	.719

a. Grouping Variable: Experimental or Control Group

The Mann-Whitney U Test results in Table 58 assess whether there is a significant difference in total verbal response scores between the 8-week control and experimental groups at the pre-test stage. The Mann-Whitney U value is 383.500, and the Wilcoxon W value is 789.500, with a Z-score of -0.360. The p-value (Asymp. Sig. (2-tailed) = .719) is well above the conventional 0.05 threshold, indicating that there is no statistically significant difference between the two groups' verbal response scores before the intervention.

These results confirm that the experimental and control groups had comparable verbal response scores at the pre-test stage, supporting the assumption that any differences observed in post-test scores would likely be due to the intervention rather than initial group differences. This finding is important as it establishes a baseline equivalence between groups, ensuring that the intervention's effects (if any) can be attributed to the program rather than pre-existing differences in verbal response abilities.

Table 60.

Ranks (Post-test, 8w Control vs 8w Experimental)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Post-test	Control Group 8 weeks	27	24.02	648.50
	Experimental Group 8 weeks	29	32.67	947.50
	Total	56		

The Mann-Whitney Ranks table (Table 60) compares the total verbal response scores between the 8-week control group and the 8-week experimental group at the post-test stage. The experimental group (M = 32.67, Sum of Ranks = 947.50) had a higher mean rank than the control group (M = 24.02, Sum of Ranks = 648.50), suggesting that participants in the experimental group performed better in verbal responses after the intervention. However, to determine whether this difference is statistically significant, we would need to examine the Mann-Whitney U test results and p-values.

Table 61.

Test Statistics^a (Post-test, 8 week comparison)

	Total Verbal Response Score of Posttest
Mann-Whitney U	270.500
Wilcoxon W	648.500
Z	-1.991
Asymp. Sig. (2-tailed)	.046

a. Grouping Variable: Experimental or Control Group

The Mann-Whitney U test results (Table 60) indicate a statistically significant difference in total verbal response scores between the 8-week experimental and control groups at the post-test stage (U = 270.500, Z = -1.991, p = .046). Since p < .05, this suggests that the experimental group outperformed the control group in verbal responses after the intervention, implying that the intervention had a meaningful impact on improving verbal response scores. The difference in ranks observed in Table 70 is now confirmed to be statistically significant, supporting the conclusion that the experimental program contributed to better verbal response performance.

Table 16a.

Ranks (Pretest, 4w Control vs 4w Experimental)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Pretest	Control Group 4 weeks	21	23.69	497.50
	Experimental Group 4 weeks	15	11.23	168.50
	Total	36		

Table 16b.

Test Statistics^a (Pretest, 4w Comparison)

Mann-Whitney U	48.500
Wilcoxon W	168.500
Z	-3.502
Asymp. Sig. (2-tailed)	<.001
Exact Sig. [2*(1-tailed Sig.)]	<.001^b

a. Grouping Variable: Experimental or Control Group
b. Not corrected for ties.

The Mann-Whitney Ranks and Test Statistics (Tables 61 and 62) indicate a significant difference in pre-test verbal response scores between the 4-week control group and the 4-week experimental group. The control group had a higher mean rank (M = 23.69, Sum of Ranks = 497.50) compared to the experimental group (M = 11.23, Sum of Ranks = 168.50). The Mann-

Whitney U value is 48.500, Z = -3.502, and p < .001, indicating that the control group started with significantly higher verbal response scores than the experimental group at the pre-test stage.

Table 64.

Ranks (Post-test, 4w Control vs 4w Experimental)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Post-test	Control Group 4 weeks	20	23.08	461.50
	Experimental Group 4 weeks	15	11.23	168.50
	Total	35		

The post-test ranks (Table 64) show a similar pattern, with the control group (M = 23.08, Sum of Ranks = 461.50) still outperforming the experimental group (M = 11.23, Sum of Ranks = 168.50). Since the ranking difference persists, this suggests that the 4-week intervention did not help the experimental group catch up to the control group in verbal response performance. The pre-existing gap between groups remained, indicating that a 4-week intervention may not have been sufficient to significantly improve verbal responses for the experimental group.

Table 65.

Test Statistics (Post-test, 4w Comparison)

	Total Verbal Response Score of Post-test
Mann-Whitney U	48.500
Wilcoxon W	168.500
Z	-3.390
Asymp. Sig. (2-tailed)	<.001
Exact Sig. [2*(1-tailed Sig.)]	<.001^{b}

a. Grouping Variable: Experimental or Control Group

b. Not corrected for ties.

The Mann-Whitney U test results (Table 65) indicate a statistically significant difference in post-test verbal response scores between the 4-week control and experimental groups (U = 48.500, Z = -3.390, p < .001). Since p < .001, this suggests that the control group significantly outperformed the experimental group in verbal responses after the intervention. Given that the pre-test scores also showed a significant gap between groups, this result implies that the 4-week intervention was not effective in closing the initial performance gap, and the experimental group did not show significant improvement relative to the control group.

Table 16c.

Test Statistics (Pretest, 8w vs 4w Experimental)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Pretest	Experimental Group 8 weeks	29	24.34	706.00
	Experimental Group 4 weeks	15	18.93	284.00
	Total	44		

Table 16d.

Test Statistics^a (Pretest, 8w vs 4w Experimental)

	Total Verbal Response Score of Pretest
Mann-Whitney U	164.000
Wilcoxon W	284.000
Z	-1.327
Asymp. Sig. (2-tailed)	.184

a. Grouping Variable: Experimental or Control Group

The Mann-Whitney Ranks and Test Statistics (Tables 16d and 16e) compare pre-test verbal response scores between the 8-week experimental group and the 4-week experimental group. The 8-week experimental group had a higher mean rank (M = 24.34, Sum of Ranks =

706.00) compared to the 4-week experimental group (M = 18.93, Sum of Ranks = 284.00). However, the Mann-Whitney U test (U = 164.000, Z = -1.327, p = .184) indicates that this difference is not statistically significant. This suggests that before the intervention, both experimental groups had comparable verbal response scores, meaning that any post-test differences would likely be due to the length of the intervention rather than initial ability differences.

Table 68.

Ranks (Post-test, 8w vs 4w Experimental)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Post-test	Experimental Group 8 weeks	29	25.67	744.50
	Experimental Group 4 weeks	15	16.37	245.50
	Total	44		

Table 69.

Test Statistics (Post-test, 8w vs 4w Experimental)

	Total Verbal Response Score of Post-test
Mann-Whitney U	125.500
Wilcoxon W	245.500
Z	-2.281
Asymp. Sig. (2-tailed)	.023

a. Grouping Variable: Experimental or Control Group

The Mann-Whitney U test results (Tables 68 and 69) compare post-test verbal response scores between the 8-week and 4-week experimental groups. The 8-week experimental group had a higher mean rank (M = 25.67, Sum of Ranks = 744.50) than the 4-week experimental group (M = 16.37, Sum of Ranks = 245.50). The Mann-Whitney U test (U = 125.500, Z = -2.281, p = .023) indicates a statistically significant difference ($p < .05$) between the two groups.

This suggests that the 8-week intervention was more effective than the 4-week intervention in improving verbal response scores, reinforcing the idea that a longer duration of intervention leads to greater improvements in verbal performance.

Table 18a.

Ranks (Pretest, 8w vs 4w Control)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Pretest	Control Group 8 weeks	28	19.09	534.50
	Control Group 4 weeks	21	32.88	690.50
	Total	49		

Table 18b.

Test Statistics^a (Pretest, 8w vs 4w Control)

	Total Verbal Response Score of Pretest
Mann-Whitney U	128.500
Wilcoxon W	534.500
Z	-3.348
Asymp. Sig. (2-tailed)	<.001

a. Grouping Variable: Experimental or Control Group

The Mann-Whitney U test results (Tables 18a and 18b) compare pre-test verbal response scores between the 8-week and 4-week control groups. The 4-week control group had a significantly higher mean rank (M = 32.88, Sum of Ranks = 690.50) compared to the 8-week control group (M = 19.09, Sum of Ranks = 534.50). The Mann-Whitney U test (U = 128.500, Z = -3.348, p < .001) indicates a statistically significant difference (p < .001) between the two groups. This suggests that, at the pre-test stage, the 4-week control group had significantly higher verbal response scores than the 8-week control group, meaning the two control groups started with different baseline abilities before any intervention.

Table 18c.

Ranks (Post-test, 8w vs 4w Control)

	Experimental or Control Group	N	Mean Rank	Sum of Ranks
Total Verbal Response Score of Post-test	Control Group 8 weeks	27	18.63	503.00
	Control Group 4 weeks	20	31.25	625.00
	Total	47		

Table 18d.

Test Statistics^a (Post-test, 8w vs 4w Control)

	Total Verbal Response Score of Post-test
Mann-Whitney U	125.000
Wilcoxon W	503.000
Z	-3.129
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Experimental or Control Group

The Mann-Whitney U test results (Tables 18c and 18d) compare post-test verbal response scores between the 8-week and 4-week control groups. The 4-week control group had a significantly higher mean rank (M = 31.25, Sum of Ranks = 625.00) compared to the 8-week control group (M = 18.63, Sum of Ranks = 503.00). The Mann-Whitney U test (U = 125.000, Z = -3.129, p = .002) indicates a statistically significant difference (p = .002) between the two groups. This suggests that even after the intervention period, the 4-week control group continued to outperform the 8-week control group in verbal response scores, highlighting that the initial pre-test differences persisted and were not altered by the length of time in the control condition.

Table 18e.

Test Statistics (Pretest and Post-test, 4w vs 8w Experimental Comparison)

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	T	Df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Total Score of Pretest	Equal variances assumed	.090	.766	.797	42	.215	.430	.828	1.039	-1.269	2.924
	Equal variances not assumed			.812	29.945	.212	.423	.828	1.019	-1.254	2.910
Total Score of Post-test	Equal variances assumed	.012	.913	2.613	42	.006	.012	2.844	1.088	.647	5.040
	Equal variances not assumed			2.695	30.962	.006	.011	2.844	1.055	.692	4.996

Table 18e presents the results of independent samples *t*-tests. For the pretest scores, the difference between the two groups was not statistically significant ($t(42) = 0.797$, $p = .430$), and the confidence interval for the mean difference (−1.269 to 2.924) includes zero. This indicates that both groups began the study with comparable baseline performance. However, for the post-test scores, the 8-week group scored significantly higher than the 4-week group ($t(42) = 2.613$, $p = .012$), with a mean difference of 2.84 points. The 95% confidence interval for this difference (.647 to 5.040) does not include zero, confirming a statistically and practically meaningful difference in favor of the longer intervention.

Section 8: Final Model Results (ANCOVA Revisited)

Table 51.

Tests of Between-Subjects Effects

Dependent Variable: Total Score of Post-test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	991.797[a]	7	141.685	198.158	<.001
Intercept	4.092	1	4.092	5.723	.019
Group	.765	3	.255	.357	.784
PretestTotalScore	802.665	1	802.665	1122.589	<.001
Group * PretestTotalScore	1.614	3	.538	.753	.524
Error	59.346	83	.715		
Total	138450.000	91			
Corrected Total	1051.143	90			

a. R Squared = .944 (Adjusted R Squared = .939)

The ANCOVA results indicate that pretest scores are the strongest predictor of post-test performance, while group membership (control vs. experimental, 4-week vs. 8-week) does not significantly impact post-test scores when controlling for pretest differences. The overall model explains 94.4% of the variance in post-test scores ($R^2 = 0.944$, $p < .001$), showing that the pretest scores are highly influential ($F = 1122.589$, $p < .001$). However, the group effect ($F = 0.357$, $p = .784$) and the Group * Pretest Interaction ($F = 0.753$, $p = .524$) were not significant, indicating that neither group membership nor an interaction between group and pretest scores meaningfully influenced post-test results.

In summary, the ANCOVA results show that children who performed well in shape recognition during the pretest continued to perform strongly on the post-test. This indicates that baseline recognition ability is a significant factor in post-test success.

Table 52.

Estimates (Adjusted Means)

Table 52 shows the estimated means for the post-test scores, adjusted for pretest scores.

Dependent Variable: Total Score of Post-test

			95% Confidence Interval	
Groups	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Control Group 8 weeks	38.120[a]	.163	37.796	38.443
Experimental Group 8 weeks	40.307[a]	.157	39.995	40.619
Control Group 4 weeks	38.224[a]	.192	37.843	38.605
Experimental Group 4 weeks	38.291[a]	.226	37.841	38.741

a. Covariates appearing in the model are evaluated at the following values: Total Score of Pretest = 37.82.

Table 52 presents the post-test means adjusted through ANCOVA, accounting for baseline (pretest) scores. The 8-week experimental group achieved the highest adjusted mean (M = 40.31, 95% CI [39.99, 40.62]), which does not overlap with the confidence intervals of any other group. This clearly demonstrates a significant learning advantage from the longer intervention. In contrast, the 4-week experimental group (M = 38.29, 95% CI [37.84, 38.74]) overlaps substantially with both the 8-week control (M = 38.12, CI [37.80, 38.44]) and 4-week control (M = 38.22, CI [37.84, 38.61]). This overlap indicates no meaningful improvement from the shorter intervention. In summary, these findings show that while the 8-week experimental approach led to significant gains over all comparison groups, the 4-week version did not outperform standard non-intervention conditions.

Table 53. Pairwise Comparisons of Adjusted Means

Dependent Variable: Total Score of Post-test

95% Confidence Interval for Difference[b]

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.^{b}	Lower Bound	Upper Bound
Control Group 8 weeks	Experimental Group 8 weeks	-2.187*	.226	<.001	-2.799	-1.576
	Control Group 4 weeks	-.105	.251	1.000	-.784	.575
	Experimental Group 4 weeks	-.171	.279	1.000	-.925	.582
Experimental Group 8 weeks	Control Group 8 weeks	2.187*	.226	<.001	1.576	2.799
	Control Group 4 weeks	2.083*	.248	<.001	1.413	2.752
	Experimental Group 4 weeks	2.016*	.275	<.001	1.272	2.760
Control Group 4 weeks	Control Group 8 weeks	.105	.251	1.000	-.575	.784
	Experimental Group 8 weeks	-2.083*	.248	<.001	-2.752	-1.413
	Experimental Group 4 weeks	-.066	.296	1.000	-.868	.735
Experimental Group 4 weeks	Control Group 8 weeks	.171	.279	1.000	-.582	.925
	Experimental Group 8 weeks	-2.016*	.275	<.001	-2.760	-1.272
	Control Group 4 weeks	.066	.296	1.000	-.735	.868

Table 53 summarizes post-hoc comparisons of adjusted post-test means (after ANCOVA), clarifying which group differences were statistically significant. The 8-week experimental group demonstrated consistently superior performance compared to all other groups: it outperformed the 8-week control by an average of 2.187 points ($p < .001$), the 4-week control by 2.083 points ($p < .001$), and the 4-week experimental group by 2.016 points ($p < .001$). By contrast, none of the other comparisons reached significance: the 4-week experimental group did not differ from either control group, and the two control groups (8-week vs. 4-week) were statistically equivalent ($p = 1.000$ in all cases).

Section 9: Change Scores & Normality

Table19a.

Descriptive Statistics for Gesture Change Scores

Statistic	Value
Mean	2.41
95% CI Lower	1.70
95% CI Upper	3.12
Median	1.00
Std. Deviation	3.41
Skewness (SE)	1.486 (.253)
Kurtosis (SE)	1.644 (.500)

Table 19b.

Tests of Normality for Gesture Change Scores

Test	Statistic	df	p-value
Kolmogorov–Smirnov	.253	91	< .001
Shapiro–Wilk	.760	91	< .001