

INTERNATIONAL STUDIES IN EDUCATIONAL SCIENCES

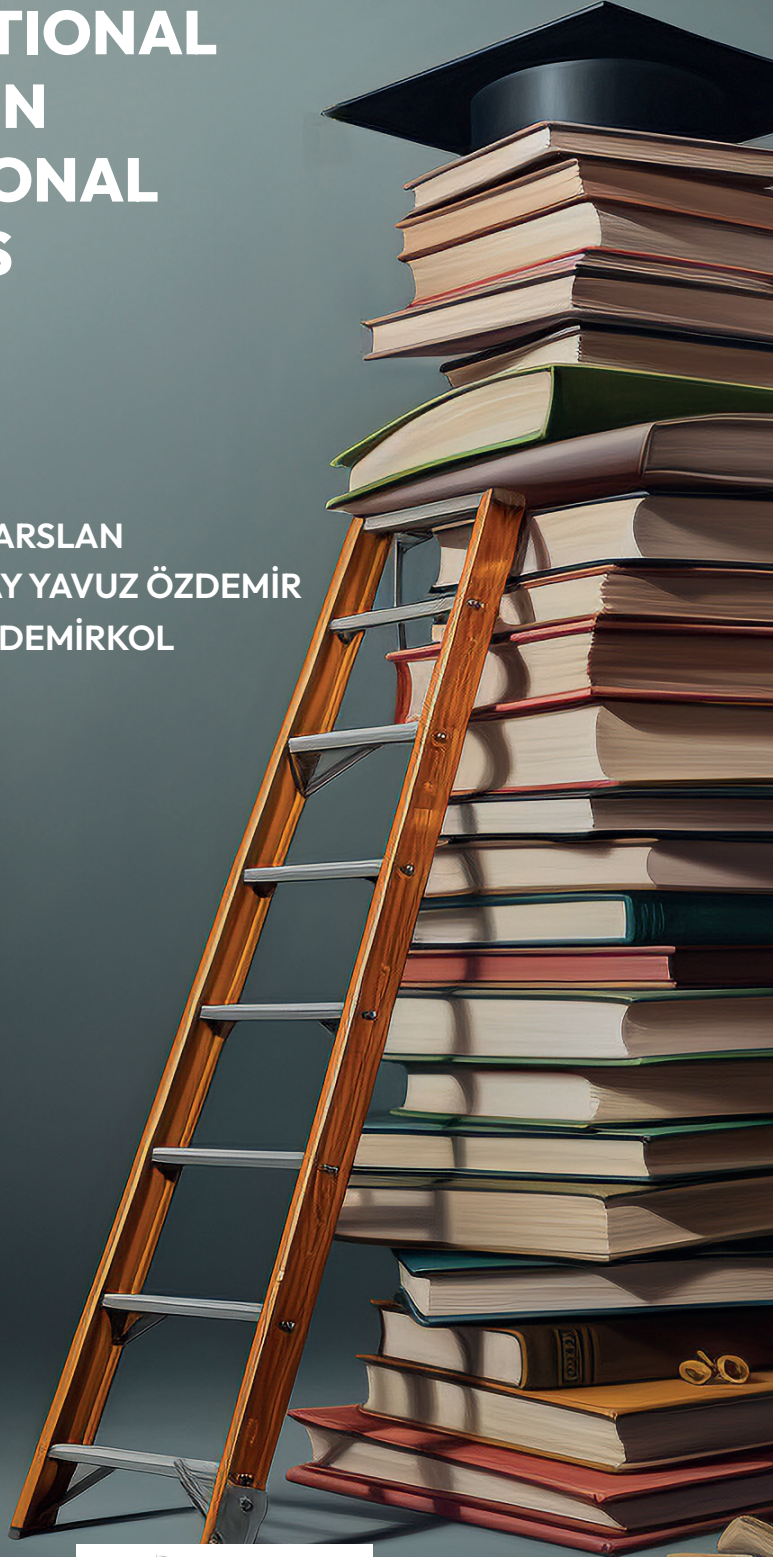
MARCH 2026

EDITORS

DOÇ. DR. AYTEN ARSLAN

PROF. DR. TUNCAY YAVUZ ÖZDEMİR

DOÇ. DR. MURAT DEMİRKOL



Genel Yayın Yönetmeni / Editor in Chief • C. Cansın Selin Temana

Kapak & İç Tasarım / Cover & Interior Design • Serüven Yayınevi

Birinci Basım / First Edition • © MART 2026

ISBN • 978-625-8671-90-2

© copyright

Bu kitabın yayın hakkı Serüven Yayınevi'ne aittir.

Kaynak gösterilmeden alıntı yapılamaz, izin almadan hiçbir yolla çoğaltılamaz. The right to publish this book belongs to Serüven Publishing. Citation can not be shown without the source, reproduced in any way without permission.

Serüven Yayınevi / Serüven Publishing

Türkiye Adres / Turkey Address: Kızılay Mah. Fevzi Çakmak 1. Sokak

Ümit Apt No: 22/A Çankaya/ANKARA

Telefon / Phone: 05437675765

web: www.seruyenyayinevi.com

e-mail: seruyenyayinevi@gmail.com

Baskı & Cilt / Printing & Volume

Sertifika / Certificate No: 47083

INTERNATIONAL STUDIES IN EDUCATIONAL SCIENCES

EDITORS

DOÇ. DR. AYTEN ARSLAN
PROF. DR. TUNCAY YAVUZ ÖZDEMİR
DOÇ. DR. MURAT DEMİRKOL

CONTENTS

CHAPTER 1

AN INTERPRETIVE READING OF CHILD DEVELOPMENT: AN EDUCATIONAL EVALUATION THROUGH THE HERO'S JOURNEY PATTERN

Aslı Ceren ALAÇAM..... 1

CHAPTER 2

SUPPORTING EXECUTIVE FUNCTIONS THROUGH SCIENCE EDUCATION IN EARLY CHILDHOOD FROM A CHILD DEVELOPMENT PERSPECTIVE

Pelin YILDIRIM..... 19

CHAPTER 3

ARTIFICIAL INTELLIGENCE BASED SOLUTION STRATEGIES IN SCIENCE EDUCATION: ANALYSIS OF HSTS QUESTIONS WITH CHATGPT

Buğra TAŞKIRAN, Mustafa ERGUN..... 33

CHAPTER 4

MAPPING WHAT RESEARCH CANNOT SAY: METHODOLOGICAL REFLECTIONS ON CROSS-NATIONAL BIBLIOMETRIC ANALYSIS OF DOCTORAL DISSERTATIONS

Bora BAŞARAN..... 57

CHAPTER 5

REFLECTIONS OF ALTERNATIVE EDUCATION EXPERIENCES ON STUDENTS' ACADEMIC PERFORMANCE AND SKILL DEVELOPMENT

Ece YILMAZ, Pınar ÇAVAŞ, Aslı Ceren ALAÇAM 79

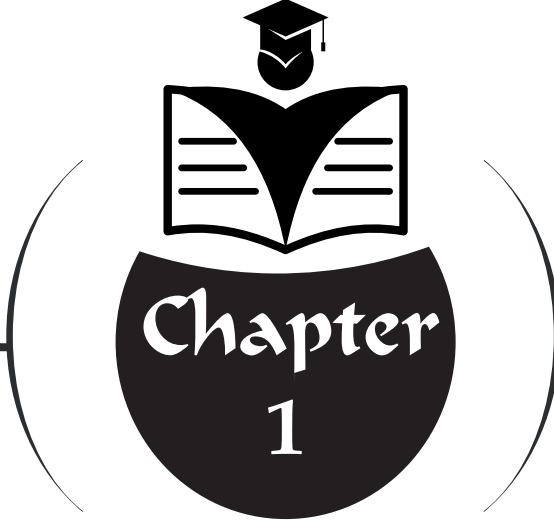
CHAPTER 6
THE IMPACT OF DIGITAL GAME-BASED LEARNING IN SOCIAL STUDIES ON THE SOCIAL-EMOTIONAL DEVELOPMENT OF GIFTED STUDENTS: AN AGE-BASED ANALYSIS

Fatih PALA101

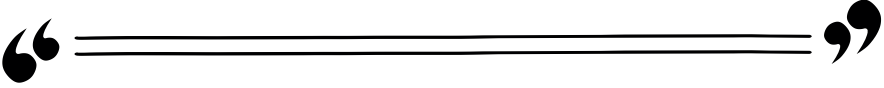
CHAPTER 7
SCENARIO BASED LEGO ROBOTICS IN SCIENCE TEACHER EDUCATION: THE CASE OF WATER THEME¹

Türkan Aybike AKARCA, Mustafa ERGUN123

¹ * This study was produced from the first author's master's thesis under the supervision of the second author.



AN INTERPRETIVE READING OF CHILD DEVELOPMENT: AN EDUCATIONAL EVALUATION THROUGH THE HERO'S JOURNEY PATTERN¹



Aslı Ceren ALAÇAM²

¹ This study is derived from the doctoral dissertation titled *Çocuk Kahramanın Yolculuğu: Çocuk Filmleri Üzerine Arketipsel Bir İnceleme* written by Aslı Ceren Alaçam under the supervision of Prof. Dr. Emet Gürel at Ege University, Institute of Social Sciences, İzmir, 2023.

² Dr., Research Assistant, Ege University, Faculty of Education, Department of Primary Education, İzmir, Türkiye, asli.ceren.alacam@ege.edu.tr, ORCID: 0000-0003-4152-0902

1. Childhood: From a Developmental Process to the Construction of Meaning

Childhood is a period in which not only biological growth but also the construction of meaning takes place. It is a formative stage during which the foundations of personality are established and which plays a decisive role in many aspects of an individual's later life. During this process, children's experiences, their relationships with themselves and with others, and the ways in which they interpret these interactions significantly shape their sense of self and their understanding of the world.

Childhood encompasses the developmental period between infancy and adulthood, including physical, cognitive, social, and emotional changes. This developmental process involves both quantitative and qualitative transformations that occur through learning, experience, and maturation (Senemoğlu, 2005). These changes form an interconnected and holistic structure. As the child's body grows, the development of the muscular, skeletal, and nervous systems occurs in coordination with one another. The physical maturation of the body influences learning processes and, consequently, psychological and social development as well (Yeşilyaprak, 2012). Therefore, growth, learning, and maturation should be understood as dynamic and interrelated processes that continuously interact with one another.

Throughout development, children pass through a series of stages. Each stage involves different challenges and requires the activation of new skills. The crises, difficulties, and inadequacies encountered during these stages become learning experiences that prepare the child for the next phase of development.

This gradual journey of growth bears a strong resemblance to the Hero's Journey, a narrative structure found in myths, folktales, and traditional stories. Across many cultures and historical periods, this narrative pattern portrays a hero who leaves their ordinary world, undergoes a process of initiation, and eventually returns transformed. Joseph Campbell (2010), who examined numerous narratives from different cultures, identified this universal structure in which the hero passes through various stages, faces trials that test specific abilities, learns through experience, and ultimately undergoes personal transformation.

From this perspective, the Hero's Journey provides a useful pedagogical metaphor for interpreting and understanding the stages of children's developmental processes.

Childhood development has long been examined by developmental and learning theorists, and these theoretical frameworks have guided educational and health-related practices. In the present study, however, the growth and developmental journey of the child is approached from a different perspective

by drawing upon mythic narratives. The pedagogical interpretations derived from the Hero's Journey pattern aim to offer a new interpretive framework for understanding the educational dimensions of child development. Such a perspective may enable teachers to evaluate children's developmental crises and learning processes from a more holistic standpoint while embracing a guiding and supportive role.

Accordingly, the aim of this study is to reinterpret child development through the framework of the Hero's Journey from an educational perspective. By drawing on theoretical foundations, the study seeks to present pedagogical insights that contribute to rethinking educational environments through a developmental lens.

2. Childhood and Development

In the literature, childhood is typically divided into several stages based on age ranges. These stages are generally classified as infancy–early childhood (0–3 years), early childhood (3–6 years), and late childhood (6–12 years), with each stage characterized by distinct developmental features.

Prior to these stages is the prenatal period, which lasts approximately nine months and prepares the human infant to adapt to life outside the womb. During this period, maternal nutrition and various environmental influences reaching the fetus affect many physiological and psychological characteristics of the baby, particularly the development of the brain (Santrock, 2016).

The infancy period, which begins at birth and continues until approximately three years of age, involves the infant's physiological adaptation to external conditions, including the regulation of body temperature, respiration, and heartbeat. During this stage, caregiving practices shape the development of feeding and sleeping patterns as well as toilet training (Yeşilyaprak, 2012). Other important developmental achievements of this period include acquiring the native language, developing symbolic thinking, and learning to walk.

Beginning around the age of two and extending until approximately six years of age, the early childhood period, which largely corresponds to the preschool years, is marked by a gradual decrease in children's dependence on adults. During this stage, children acquire self-care skills, while their language and cognitive development accelerate. They also begin to learn gender differences, develop social relationships, and gradually form an understanding of moral rules. Although children's behavior remains largely egocentric, they gradually start to recognize perspectives other than their own (Bayhan & Artan, 2004).

The period between six and eleven years of age, during which children typically attend primary school, is referred to as late childhood. During this

stage, physical growth continues and motor skills develop further. At the same time, children begin to experience social and emotional challenges through their interactions and experiences, which contribute significantly to the development of their sense of self (Santrock, 2016).

The period between 11 and 18 years of age, often referred to as adolescence, represents the transitional stage before adulthood. This period is characterized by rapid physical and cognitive changes. Adolescents begin to form their own values and life goals, seek greater independence from their families, and reflect on their roles and responsibilities within society. They also begin to internalize gender roles while planning their future (Yeşilyaprak, 2012).

Throughout this developmental journey, children acquire competencies in three major domains: physical, cognitive, and psychosocial development. These domains develop through interrelated processes that interact with and support one another.

Physical and Psychomotor Development

Although children's physical development generally follows a predictable sequence, the pace and extent of this development may vary depending on genetic factors, nutrition, and cultural conditions (Tepeli, 2012).

Motor development refers to the increasing ability of individuals to control their bodies. In other words, it involves the development of behaviors based on coordination between the mind and muscles. For motor development to occur, sensory organs, muscles, and the nervous system must function in coordination (Senemoğlu, 2005).

Physical growth and the corresponding development of motor skills play an essential role in children's adaptation to the world and later facilitate their participation in social life.

Cognitive Development

Cognitive development encompasses the internal mental processes involved in functions such as attention, perception, memory, language development, reading and writing, problem-solving, reasoning, and creativity.

According to Piaget, cognitive development emerges through the interaction of hereditary characteristics and environmental experiences. This process involves not only the maturation of the brain and nervous system but also the integration of experiences that help the individual adapt to the environment (Bayhan & Artan, 2004). Piaget argued that when children encounter new objects, events, or situations, their cognitive equilibrium is disrupted. In response to this imbalance, children attempt to restore equilibrium by adapting their cognitive structures. This process

of equilibration occurs with each new experience and leads to learning and development (Senemoğlu, 2005).

In this context, Piaget (1971) proposed that knowledge is constructed through the learner's experiences and actions. This idea later formed one of the theoretical foundations of the constructivist approach in education, which emphasizes learning through active engagement and experience (as cited in Trawick-Smith, 2014).

Another theorist who emphasized the role of experience and social interaction in cognitive development and learning was Jerome Bruner. According to Bruner, learning is "an active process in which learners construct new ideas or concepts based upon their current or past knowledge" (Arslan, 2012, p. 19). Within this active process, children's ability to explore and discover independently plays a crucial role.

Similarly, Lev Vygotsky highlighted the importance of the social environment in learning and cognitive development. According to Vygotsky, children acquire concepts and ideas through their interactions with their environment, gradually developing skills and attitudes through these experiences. The cultural environment in which the child grows up therefore plays a highly influential role. Through interactions with adults, children not only acquire knowledge but also begin to construct their own psychological processes and self-understanding (Senemoğlu, 2005).

Vygotsky's emphasis on the role of adults, the social context, and collaboration has also guided educational practices. In classroom settings, teachers act as adult guides who support learning, while peers contribute to learning through cooperation (Berk, 2013). Over time, children gradually develop self-regulation and become capable of directing their own learning processes. Therefore, educational environments in which teachers assume a guiding role while children take responsibility for their own learning are of great importance.

Later developments based on Piaget's theory have further emphasized information processing, cognitive strategies, and the detailed structure of cognitive stages. Within this perspective, abilities such as critical thinking, creative thinking, scientific reasoning, and problem-solving have become important indicators of cognitive development.

From an educational perspective, one of the most significant implications of cognitive theories is the role of teachers and other adults in guiding children toward increasingly complex ways of thinking. Through social interaction, children first experience knowledge within a social context and gradually internalize it as part of their individual cognitive structures. Knowledge initially emerges within social interaction is reorganized at the individual

level and transformed into more complex cognitive structures. In this way, cognitive development progresses gradually through interaction and guided learning (Senemoğlu, 2005).

Emotional and Social Development

Changes in individuals' emotions, interpersonal relationships, and personality constitute the processes of socio-emotional development. In the literature, these processes are sometimes examined separately and sometimes within the broader frameworks of moral development and personality development (Koç et al., 2001; Bacanlı, 2003; Yeşilyaprak, 2012).

Emotions, which represent the outward expression of individuals' subjective experiences, are fundamental components of their inner world (Erden, 2012). The development of these emotional capacities is shaped by both biological and environmental factors. Infants' ability to regulate their emotions is closely related to the development of specific regions of the brain. At the same time, social relationships contribute significantly to the diversification and development of emotional expressions. For this reason, emotional expressions and emotional development may vary across cultures. The level of emotional maturity individuals achieve during development influences how they respond to experiences and maintain emotional balance in later life.

Social development, on the other hand, can be defined as the dynamic process through which individuals acquire the knowledge, skills, and values necessary to establish effective relationships with others and to adapt to society (APSMHI, 2013, cited in Yurtsever Kılıçgün, 2015, p. 111). The quality of individuals' interactions with others plays a crucial role in the development of their sense of self and their ability to form healthy relationships (Turan & Yükselen, 2021).

Socialization begins with interactions within the family from birth and gradually expands through relationships with the immediate environment, peers, and school settings. Schools, in particular, play an important role in this developmental process (Berk, 2013). Through social interaction, individuals also become aware of the gender roles expected by society and gradually shape their identities in relation to these expectations (Orçan, 2012).

Albert Bandura focused on the role of social interaction in learning and developed Social Cognitive Theory. According to this theory, individuals acquire a wide range of behaviors, thoughts, and emotions by observing others in their environment. These observational processes continue throughout life and contribute to continuous development (Bandura, 1977; 1986). Children acquire many skills through modeling. The rewards and punishments associated with observed behaviors influence their choices and contribute to the development of self-efficacy beliefs (Bandura, 1999).

For learning and development to progress in a healthy manner, the environments in which children grow, including family and school contexts, should support the development of children's self-efficacy beliefs and self-regulation skills. Bandura's Social Learning Theory and its underlying principles have significantly influenced educational practices by emphasizing the importance of role models such as parents, teachers, peers, and other adults in children's developmental environments.

Taken together, these theoretical perspectives suggest that childhood development is not merely a process of biological growth. Rather, it is a dynamic journey in which individuals encounter challenges, interact with their environment, learn through experience, and gradually undergo transformation. In this sense, childhood development can also be interpreted as a developmental journey through which individuals mature as they pass through different stages. The Hero's Journey pattern found in myths and narratives provides one of the most powerful metaphorical frameworks for interpreting this process of transformation.

3. Interpreting Child Development Through the Pattern of the Hero's Journey

Rank (2016) argued that the process of individual development is closely related to the recurring hero narratives found in myths. According to him, the growing individual's attempt to break away from parental authority may be understood as a form of the "hero myth." In this sense, the ego of the mythical hero resembles the ego of the child. Likewise, themes such as the stepmother or abandonment, which appear in myths, correspond to themes found in childish and neurotic fantasies, thereby supporting this argument. This perspective suggests that the Hero's Journey is not merely a mythological structure but may also be read as a symbolic representation of human development.

Developmental processes in human life are not limited to biological and cognitive changes; they also involve processes of transformation through which individuals make sense of their experiences. In this respect, childhood is a dynamic period during which individuals acquire new skills, encounter various developmental thresholds, and continuously reconstruct themselves through new experiences. A similar transformation can be observed in the Hero's Journey pattern found in myths and narratives. The hero's departure from the familiar world, passage through various trials, and eventual return to the point of origin after transformation symbolically represent the human experience.

From this perspective, childhood may also be interpreted as a journey shaped by developmental tasks, crises, and learning experiences. In this section, the stages of the Hero's Journey will be examined in relation to child development, and their possible counterparts in childhood will be discussed within a theoretical framework.

In Campbell's monomyth theory, the Hero's Journey consists of three major stages, Departure (Separation), Initiation, and Return, along with their sub-stages (Campbell, 2010). The well-known screenwriter and author Christopher Vogler (1992/2009), adapting Campbell's 17-stage model to the context of film and television screenplays, reformulated it into a more concise structure. In Vogler's framework, these stages are identified as the Ordinary World, Call to Adventure, Refusal of the Call, Meeting with the Mentor, Crossing the Threshold, Tests, Allies and Enemies, Approach to the Inmost Cave, The Ordeal, Reward, The Road Back, The Resurrection, and Return with the Elixir.

In this part of the study, the correspondences between the stages of the Hero's Journey described by Campbell (2010) and Vogler (2009) and the developmental period of childhood will be examined. Childhood can be understood as a journey as a whole, but important experiences within childhood may also be seen as distinct cycles of journeying in themselves. Accordingly, the parallels with the Hero's Journey will be discussed both at the level of the broader developmental process and through specific examples drawn from childhood experiences.

Nearly all fairy tales and mythic narratives begin by depicting the protagonist within the Ordinary World. At this first stage of the Hero's Journey, the hero lives comfortably within the familiar conditions of that world (Campbell, 2010). In the context of the child's life journey, the first representation of the ordinary and comfortable world may be interpreted as the period in the mother's womb. From conception to birth, the fetus remains in a secure and protected environment sustained by the mother's bodily resources. During this process, it is nourished through the mother and indirectly affected by her psychological state as well as by the environmental conditions influencing her. In this sense, the infant occupies a passive position within this zone of comfort.

The moment of birth may therefore be interpreted as the beginning of the infant's adventure in the world, since it requires leaving a zone of comfort and adapting to entirely new conditions. In this respect, birth can be understood as the Call to Adventure for the human infant. In narrative structures, the hero's departure fundamentally signifies leaving behind an accustomed state of existence. The Call to Adventure, as the first step out of the comfort zone, may be external or internal, physical or psychological. Refusing such an instinctive call may mean suppressing one's impulses, which can ultimately lead to unhappiness (Dökmen, 1983). For infants, one generally cannot speak of a conscious refusal of the call; however, difficult births, post-delivery placental complications, or breech presentation may be interpreted as symbolic forms of the Refusal of the Call. Whether conscious or unconscious, such refusal still belongs to the logic of the journey. For the journey to begin, however, the call must eventually be accepted (Sarıççek, 2020).

In order to adapt to the world and continue growing, the child must accomplish a series of developmental tasks. Each of these tasks may be interpreted as an adventure in itself. Skills such as crawling, walking, grasping objects, or learning to eat are all acquired gradually through repeated attempts. Piaget's theory (1971) suggests that knowledge is constructed through the learner's actions. Through action, the child constructs knowledge about both the external world and the self. During the 0–3 age period, infants' knowledge is largely based on intuition and action. At this stage, intelligence may be understood as the capacity to satisfy needs through movement and perception (as cited in Trawick-Smith, 2014). In this sense, accepting the call to adventure and progressing along the journey may be seen both as an indicator of intelligence and as a factor that contributes to its development.

Another crucial stage of the Hero's Journey is the Meeting with the Mentor, which is especially important in the context of children's lives because children, unlike adults, have not yet developed sufficient competence to act entirely on their own. In mythology and folklore, mentors serve to protect, educate, test, train, guide, and provide magical gifts to the hero (Vogler, 2009, p. 179). For a child between the ages of 0 and 3, this guiding figure is primarily the caregiver, that is, the parents. Mentors are not merely providers of material support; they also encourage the hero to overcome fear and embark on the adventure. In this regard, the role of parents at this stage becomes significant when they encourage children to learn independently, to begin walking, and to gain confidence through supportive gestures.

In mythological narratives, the guides encountered by the hero may be extraordinary beings possessing exceptional powers (Campbell, 2010; Vogler, 2009). From the child's perspective, parents and teachers may likewise appear as figures endowed with extraordinary, almost superhuman abilities. Adults are physically larger, more capable, and quicker than children. They can skillfully use tools that children cannot yet handle. When adults perform with ease tasks that children find difficult, they may appear fascinating and even magical to them. In this sense, parents and teachers closely correspond to the mentor figure in the Hero's Journey.

Guides may also create conflict or provide misguided direction within a story. In children's lives, such situations can be observed in child–parent conflicts or in forms of parenting that fail to align with sound pedagogical principles. Numerous studies show that inappropriate parental attitudes may negatively affect children's learning and development (Berk, 2013; Santrock, 2016; Trawick-Smith, 2014). Yet, just as in heroic narratives, such negative experiences in childhood may later become opportunities for growth and transformation.

The mentor's facilitating role helps the hero move to the next stage, namely Crossing the First Threshold. In narrative structures, the threshold marks the boundary between the hero's ordinary world and a different, special realm. In childhood, crossing the threshold may be interpreted as accomplishing a developmental task for example, learning to walk. The first threshold represents the child's entry into a special world. Once the child begins to walk, the range of what can be seen, reached, and explored expands. This newly accessible realm also becomes a space in which new knowledge and skills can be acquired.

However, entry into this special world may involve encounters with Threshold Guardians. In narratives, Threshold Guardians emerge at various points to block the path and test the hero, especially around entrances and transitional passages (Vogler, 2009, p. 193). In infancy, adult figures such as parents and teachers may take on this role in addition to serving as guides. For example, during the walking stage, parents may prevent or supervise the child's passage into spaces they perceive as dangerous. Vogler (2009) suggests that in encounters with Threshold Guardians, different strategies are possible—ignoring them, confronting them with faith, or negotiating with them—and that, when approached appropriately, such figures may even become allies. In this sense, parents can function both as guardians of the threshold and as allies.

In Campbell's monomyth, after crossing the first threshold the hero enters the stage known as the Belly of the Whale. This stage symbolically represents a process in which the hero undergoes transformation and prepares for rebirth before entering the unknown world. Evoking the image of the womb, the Belly of the Whale signifies a space in which the hero turns inward to confront darkness and the unknown; in this process, the previous mode of existence is symbolically destroyed and a new self emerges (Campbell, 2010, pp. 107–111). Psychologically, this dark space may be interpreted as a metaphor for confronting the unconscious and for the reorganization of the self.

In early childhood (0–3 years), it is difficult to observe such internal transformations directly. Since cognitive and linguistic abilities are still developing, these processes can only be assessed indirectly through behavioral changes. Developmental psychology literature similarly emphasizes that in early childhood, inner regulation and the formation of the self are often interpreted through behavioral signs (Erikson, 1963; Piaget, 1952). In this sense, periods during which the child turns inward and then re-emerges with new behavioral patterns may symbolically be associated with the Belly of the Whale.

In myths and fairy tales, once the hero has crossed the threshold, they enter the special realm and begin the process of Initiation. Campbell describes

this stage as one in which the hero moves through a dreamlike world of shifting and ambiguous forms and passes through a series of trials (Campbell, 2010, p. 113). Campbell's "Road of Trials" corresponds to Vogler's Tests, Allies and Enemies. This realm, previously unknown to the hero, may seem new and frightening, yet it is also the place where development, learning, and skill acquisition occur. Mythic narratives typically portray this stage as a dynamic sequence of miraculous tests and ordeals. The hero must rely on previously acquired tools, resources, and experience in order to proceed.

For a young child, this stage may be interpreted as the new experiential field entered after learning to walk. Once walking begins, the child starts exploring a world full of uncertainty—touching, tasting, and observing things never encountered before. For the child, this world may be magical, surprising, or frightening, much like the worlds found in myths and fairy tales. Every curb, step, or dark corner may become a possible enemy or a site of testing. Allies in this world may include parents, playmates, or even a comforting toy. At such moments, what the child has learned from caregivers—or sometimes simply a hand extended by a parent—may provide support.

After Tests, Allies and Enemies, Vogler (2009) identifies the next stage as the Approach to the Inmost Cave, when the hero reaches the boundary of the desired but dangerous object or space. The inmost cave is mysterious and threatening. Upon reaching this threshold, the hero may need to pause, prepare, make plans, and find a way past the guards. By this point, the hero has already gained certain skills, resources, and companions from earlier experiences in the special world, and these can now be used in approaching the dangerous zone.

This stage may be interpreted in childhood as a newly walking child approaching the other rooms of the house in order to explore them. By this stage, the child has already developed walking ability. To reach the distant, dark, and unknown room without a parent, the child needs courage. At the same time, the child must have the mother's permission to enter, because if she refuses, she may simply pick the child up and carry them away. As a hero, the child may attempt to persuade the mother not only through physical resistance but also through speech or crying. These efforts depend on which strategies the child has already learned in order to influence the caregiver.

According to Bowlby's Attachment Theory, the stage between approximately 6–8 months and 18 months–2 years is known as the clear-cut attachment phase, during which attachment between the child and the familiar caregiver becomes established. During this phase, children may show separation anxiety or protest when separated from the caregiver. Between 18 months and 2 years, children also begin to develop strategies of persuasion and negotiation in order to achieve what they want (Berk, 2013). From the

perspective of the Hero's Journey, these behaviors may be interpreted as part of the child's preparation for navigating the rules of the newly discovered "special world" entered through walking. In this sense, the competencies developed within the attachment relationship may help the child pass through the stage of the Approach to the Inmost Cave.

In Campbell's (2010) account of initiation, unlike Vogler's simplified model, the stage of Meeting with the Goddess also appears. This stage refers to the hero's encounter with various manifestations of the goddess archetype. Campbell states that the mature hero ultimately longs to return to the image of the "comforting, nourishing, young, beautiful, and good mother" retained in childhood memory. Yet the mother may not always appear in this idealized form.

Campbell (2010) describes the negative mother figure in narrative terms as follows:

The "bad" mother may appear as the object of aggressive fantasies and the feared source of retaliation; as the absent, inaccessible mother; as the prohibiting, punishing, restraining mother; as the mother who tries to keep the growing child from separating; or as the desired but forbidden mother whose presence becomes a trap for dangerous desire (p. 127).

The maternal figures Campbell identifies in myths and fairy tales can also be interpreted as corresponding to the types of mothers encountered by children in real life during the 0–3 age period. At times, the same mother may display different versions of the maternal figure. Freud's psychosexual theory similarly argues that the child's early relationship with the mother has lasting consequences for adult psychology. In the Oral Stage (0–18 months), the extent to which the child's needs are met and the mother's attitudes during this period may shape later issues of dependency and trust (Erden & Akman, 2021). Likewise, in the Anal Stage (18 months–3 years), the mother–child relationship plays a crucial role in the development of autonomy (İnci, 2012).

In the Hero's Journey, there are also female figures who guide the hero and help break the hero's chains. These figures may be interpreted, in terms of child development, as reflections of the "mother" who enables the child to achieve autonomy through appropriate attitudes and support. From this perspective, the different versions of the maternal archetype may serve as a metaphorical tool for understanding the diverse effects of maternal figures on child psychology.

In Vogler's (2009) model, the Ordeal is one of the turning points of the story and refers to the stage at which the hero faces the most formidable opposition. At this stage, the hero is tested through severe difficulties and often experiences a major loss or breakdown. Vogler describes the essence

of this process with the phrase: “Heroes must die so that they can be reborn” (2009, p. 223). In this sense, the ordeal is a critical threshold at which the hero’s old self is shaken and transformation becomes possible. Confronting death expands the hero’s experience and makes possible a deeper awareness and wisdom.

In childhood, this stage may be interpreted through the physical difficulties children encounter in their attempts to explore the world—falling, getting hurt, or suffering minor injuries. After such experiences, the child is no longer exactly the same as before. The child comes to know the possible consequences of actions and gains awareness that one can survive difficulty. In this sense, the child’s new consciousness may be interpreted as a symbolic form of rebirth.

After enduring the ordeal, the hero reaches the stage of Reward. In Campbell’s account, the final reward may take the form of a gift, a new ability, or even marriage. Psychologically, however, the reward may be understood as the transformation itself—the result of overcoming fear and confronting one’s shadow.

In the child’s life, this stage may take the form of receiving gifts, praise, or applause in response to a correct or courageous action. Kohlberg argued that in the pre-conventional level of moral development, when children have not yet internalized moral principles, they tend to evaluate right and wrong in terms of reward and punishment and according to the satisfaction of their immediate needs (Erden & Akman, 2021). From this perspective, for children aged 0–3, notions of good and bad are strongly shaped by reward and punishment. Thus, when the reward stage is considered in terms of what parents provide, it acquires pedagogical significance. At the same time, the knowledge, confidence, and courage children gain through their acts of curious exploration may also be understood as their reward. In this respect, “reward” plays a significant role in the child’s life, just as it does in the hero’s life.

The next stage, The Road Back, refers to the point at which the hero, having passed through the ordeal and gained the reward, must decide whether to remain in the special world or return to the ordinary world. The return is important because it allows the hero to carry the knowledge and experience gained in the special world back into the ordinary world from which the journey began (Vogler, 2009).

A similar dynamic may be observed in childhood. When children encounter new experiences and spaces of exploration, they may at times become captivated by the “special world” and not wish to return, or alternatively, they may move toward return out of anxiety. At this point, the child’s level of self-efficacy may become decisive in navigating between the

special and ordinary worlds. Self-efficacy is defined as the individual's belief in their capacity to organize and successfully perform a specific action in a given situation (Bandura, 1981, cited in Koç et al., 2001, p. 155). If, through the journey, the child develops a sense of self-efficacy, the process of deciding to return may become easier. Believing that one can return-and embark on a new adventure again-may motivate the child to continue.

Another stage within the return sequence is Resurrection, which Vogler (2009) describes as one of the most critical moments both for the hero and for the narrative itself. At this stage, before returning to the Ordinary World, the hero must gain one final quality. The struggle at this point should reveal both the strengths of the hero's former self and the new capacities acquired during the journey. In childhood, learning occurs through experience, and the child symbolically reconstructs the self by adding new knowledge and skills to previous ones. In this respect, the child's developmental process may be seen as parallel to the stage of Resurrection.

Finally, Vogler (2009) states that a true hero returns to the ordinary world with something to share. In the stage of Return with the Elixir, the hero brings back a gift, and through this gift the ordinary world becomes better, healthier, and more whole. While the reward is something the hero gains personally, the elixir is something brought back for the community. In the case of the child hero, this "community" may be interpreted as the parents. With every adventure, the child becomes more competent through newly acquired knowledge and skills, and this may increase the comfort of the parents as well. At the same time, what has been learned also makes future adventures possible, thereby sustaining a continuous cycle of learning and development.

4. Conclusion and Educational Implications

Like the narrative of the Hero's Journey, human development may also be understood as an ongoing process of transformation. Much like a mythical hero, the individual in everyday life moves from one adventure to another throughout the course of life. Each adventure leads to new learning, and learning itself never truly comes to an end. The flow of life continually calls the individual toward new experiences and new stages of development. This is especially evident in childhood, when such "adventures" occur frequently and are often repeated until learning takes place.

In this study, childhood was examined within the framework of developmental psychology literature and interpreted through the stages of the Hero's Journey. The analysis revealed a number of parallels between the child's developmental journey and the stages of the Hero's Journey. When these literature-supported similarities are taken into consideration, the Hero's Journey pattern appears to provide a powerful metaphorical framework for understanding child development. In this respect, the study supports the view

that the developmental process of childhood can be interpreted through the motif of the Hero's Journey.

This perspective also has important implications for educational settings. Children's learning and developmental processes, much like the Hero's Journey itself, may be understood as gradual and cyclical processes involving experience, exploration, error, and transformation. From this standpoint, it may be beneficial to design educational environments as developmental journeys. Such environments should create learning opportunities by allowing children to experiment, explore, make mistakes, and try again.

Within learning environments grounded in this idea of a developmental journey, the role of the teacher can also be reconsidered. In a manner similar to the mentor figure in the Hero's Journey, teachers may function as guides who encourage children and help them acquire the knowledge and skills needed for the next stage of development. At times, this guide may take the form of a compassionate rescuer who reaches out to support the child in difficult moments; at other times, the teacher may assume a role closer to that of a threshold guardian, challenging the child, setting limits, and helping them confront themselves. Both roles can contribute meaningfully to the child's developmental process.

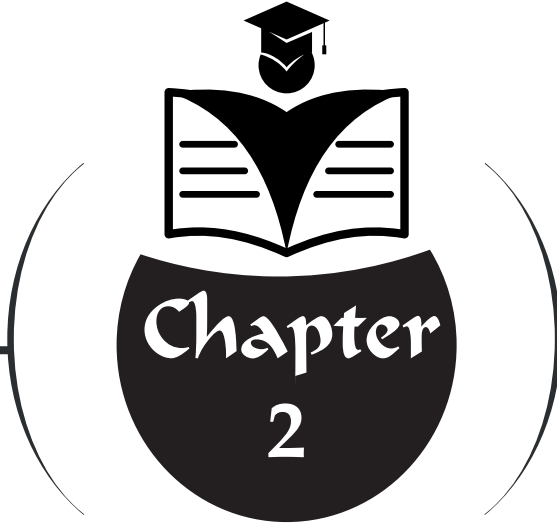
The most developmentally supportive form of guidance, however, is one that nurtures children's curiosity, encourages their desire to explore, and provides them with a safe space for learning. Curiosity and the desire to explore are among the most important forces that move the hero toward the next stage of the journey. In the same way, they are central to children's developmental progress. When educational environments offer secure opportunities for growth, children may find the inner strength to continue their journeys even when they make mistakes or encounter difficulties. In this sense, framing development as a heroic journey may also support children's self-confidence and motivation.

Ultimately, the Hero's Journey pattern should not be regarded as a framework intended to replace established theoretical models of child development. Rather, it may be understood as a metaphorical lens that helps interpret these developmental processes from a different perspective. Viewing children's developmental experiences through the metaphor of a journey may contribute to a more holistic understanding of their processes of transformation. At the same time, this perspective may also support the rethinking and redesign of educational environments in ways that more fully respond to children's developmental needs.

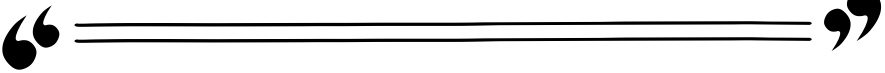
REFERENCES

- Arslan, E. (2012). Erken çocuklukta bilişsel gelişim. In M. E. Deniz (Ed.), *Erken çocukluk döneminde gelişim* (pp. 1–24). Ankara: Maya Akademi.
- Bacanlı, H. (2003). *Gelişim ve öğrenme* (6th ed.). Ankara: Nobel.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1999). Social cognitive theory: An agentic perspective. *Asian Journal of Social Psychology*, 2(1), 21–41.
- Bayhan, P. S. ve Artan, İ. (2004). *Çocuk gelişimi ve eğitimi*. İstanbul: Morpa.
- Berk, L. E. (2013). *Çocuk gelişimi* (B. Onur & A. Dönmez, Trans.). Ankara: İmge
- Campbell, J. (2010). *Kahramanın sonsuz yolculuğu* (S. Gürses, Trans.). İstanbul: Kalcı.
- Dökmen, Ü. (1983). Pinokyo'nun arketipler ve ana baba çocuk ilişkileri açısından incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 16, 2–10.
- Erden, M., & Akman, Y. (2021). *Eğitim psikolojisi* (25th ed.). Ankara: Arkadaş.
- Erden, Ş. (2012). Duygusal gelişim. In M. E. Deniz (Ed.), *Erken çocukluk döneminde gelişim* (4th ed., pp. 183–217). Ankara: Maya Akademi.
- Erikson, E. H. (1963). *Childhood and society*. New York: Norton.
- Koç, M., Yavuzer, Y., Demir, Z., & Çalışkan, M. (2002). *Gelişim ve öğrenme*. Ankara: Atlas.
- Orçan, M. (2012). Sosyal gelişim. In M. E. Deniz (Ed.), *Erken çocukluk döneminde gelişim* (4th ed., pp. 127–178). Ankara: Maya Akademi.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Rank, O. (2016). *Kahramanın doğuş miti* (G. Yavaş, Trans.). İstanbul: Pinhan.
- Santrock, J. W. (2016). *Yaşam boyu gelişim* (G. Yüksel, Ed.). Ankara: Nobel.
- Sarıççek, M. (2020). *Modern kahramanın mitolojik yolculuğu*. İstanbul: Ötügen.
- Senemoğlu, N. (2005). *Gelişim, öğrenme ve öğretim: Kuramdan uygulamaya*. Ankara: Pegem Akademi.
- Tepeli, K. (2012). Motor (hareket) gelişim. In M. E. Deniz (Ed.), *Erken çocukluk döneminde gelişim* (4th ed., pp. 91–124). Ankara: Maya Akademi.
- Trawick-Smith, J. (2014). *Erken çocukluk döneminde gelişim* (B. Akman, Trans. Ed.). Ankara: Nobel.
- Turan, F. ve Yükselen, A. İ. (2021). *Çocuk gelişimi I*. Ankara: Hedef

- Vogler, C. (2009). *Yazarın yolculuđu: Senaryo ve öykü yazımının sırları*. İstanbul: Okuyan.
- Yurtsever Kılıçgün, M. (2015). Sosyal gelişim. In E. Arslan (Ed.), *Erken çocukluk döneminde gelişim* (pp. 111–129). Ankara: Eğiten Kitap.



**SUPPORTING EXECUTIVE
FUNCTIONS THROUGH
SCIENCE EDUCATION IN
EARLY CHILDHOOD FROM
A CHILD DEVELOPMENT
PERSPECTIVE**



Pelin YILDIRIM¹

¹ Dr., Firat University, Faculty of Education, Department of Mathematics and Science Education, Elazig, Turkey, 0000-0003-4425-2472

1. Introduction

Early childhood (ages 0-6) is regarded as a critical developmental stage in the field of child development, during which the foundations of cognitive, linguistic, social, and emotional development are established and learning occurs in the most rapid and lasting manner (Diamond, 2013; Lezak, 1995). During this period, children make sense of the world through their interactions with their environment, develop concepts based on their experiences, and learn to use language not only as a means of communication but also as a fundamental cognitive tool that structures thinking, problem solving, and learning (Miyake et al., 2000).

Early childhood is a critical life stage during which individuals' cognitive, emotional, and social development occurs rapidly. The fundamental skills acquired during this period directly influence individuals' success in their subsequent educational experiences and daily lives. In particular, the cognitive processes referred to as executive functions play a central role in children's ability to plan, sequence, and regulate goal-directed behaviors. Executive functions consist of core components such as working memory, inhibitory control (inhibition), and cognitive flexibility, and they support individuals' problem-solving, planning, reasoning, and learning processes (Diamond, 2013; Lezak, 1995; Miyake et al., 2000). These functions are critical not only in cognitive tasks but also in coping with social and emotional situations.

Science education is a learning domain that supports children's natural curiosity and motivation to learn, while fostering problem-solving and scientific thinking skills. Early exposure to science education enables children to acquire conceptual and procedural knowledge, prepares them cognitively and socially for primary school, and helps them develop a science learner identity (Brenneman et al., 2009; Clements & Sarama, 2016; Oppermann et al., 2018). This process allows children to develop fundamental scientific process skills such as exploring the environment, making observations, formulating hypotheses, and testing them.

The relationship between executive functions and science education is particularly important during early childhood. In science activities, children use their executive functions in processes such as generating hypotheses, making predictions, and evaluating outcomes, and these processes contribute to the development of their cognitive control mechanisms (Gropen et al., 2011; Zelazo et al., 2006). Moreover, research indicates that executive functions can be strengthened not only through biological maturation but also through appropriate pedagogical approaches (Blair, 2016; Diamond et al., 2007). Therefore, early childhood science education should be structured in ways that support children's executive functions, such as attention, working memory, and inhibitory control.

In this book chapter, first, the definition, components, and neural basis of executive functions will be addressed; subsequently, the theoretical foundations of science education in early childhood will be examined; and finally, the interaction between science education and executive functions will be evaluated, and implications for educational practice will be presented. In this way, the relationship between the role of executive functions in early childhood science education design and pedagogical strategies will be systematically articulated.

2. Executive Functions: Definition, Components, and Neural Basis

Executive functions are defined as an umbrella term referring to a set of interrelated cognitive processes that enable individuals to plan, sequence, and regulate goal-directed behaviors (Anderson, 2002). These functions encompass abilities such as temporarily holding information in memory, inhibiting automatized responses, and flexibly shifting attention between relevant aspects of a task or problem (Blair, 2017). In other words, executive functions consist of complex mental skills that make it possible to set goals, plan, initiate actions, and achieve goals despite interruptions. Planning, problem solving, reasoning, and the execution of daily life skills are among the core indicators of executive functions (Blair, 2017).

The reason these functions are referred to as “executive” lies in their supervisory role in integrating information stored in different regions of the brain and their potential to influence other domains of cognition — such as learning, memory, language, and visual perception (Salimpoor & Desrocher, 2006). Executive functions enable the inhibition of habitual behaviors, the deliberate focusing of attention, and the regulation of thoughts in complex or stressful situations (Blair, 2017). They have been examined in research conducted within the fields of cognitive psychology, neuropsychology, developmental psychology, and educational sciences, leading to the development of various definitions across disciplines (Anderson, 2012; Gartstein et al., 2013; Groppe & Elsner, 2015; Hendry et al., 2016; Lezak, 1995; Miyake et al., 2000).

An examination of different definitions reveals that (Lezak, 1995) conceptualized executive functions as the capacity to successfully carry out independent, goal-directed behavior and subsumed them under intellectual, verbal, perceptual, memory, conceptual, orientation, and attentional functions. Elliott (2003) defined executive functions as an umbrella term for cognitive processes such as solving novel problems, modifying behavior in light of new information, developing strategies, and sequencing complex actions (Diamond, 2013); described these functions as top-down mental processes required in situations where automatic or intuition-based responses are insufficient. Blair (2017), in turn, conceptualized executive functions as

thinking skills that support reasoning, planning, problem solving, and the management of one's life.

Although there is no single, universally accepted definition of executive functions, there is a general consensus among researchers regarding three core components: Working memory, inhibitory control (inhibition), and cognitive flexibility (Diamond, 2013; Miyake et al., 2000; 18). These core components give rise to higher-order executive functions such as reasoning, problem solving, and planning (Collins & Koechlin, 2012; Lunt et al., 2012) and are critical for both normal mental functioning and social and psychological development (Corcoran & O'Flaherty, 2017).

Jacobson and colleagues (2017) classified executive functions into lower- and higher-level skills. Lower-level skills include basic functions such as cognitive fluency and processing speed, whereas higher-level skills encompass cognitive flexibility, working memory, inhibitory control, planning, reasoning, and problem-solving abilities. Korkmaz and colleagues (2013) conceptualized executive functions as “memory”, “attention”, “volition and behavior”, and “thoughts”, and discussed the indicators of each component. Memory was defined as strategies for controlling and encoding information; attention as the ability to focus, sustain, and disengage attention; volition and behavior as goal setting and the execution of complex behaviors; and thought as the abilities to understand danger, anticipate outcomes, and classify information. Gioia and colleagues (2000) emphasized that executive functions are not limited to cognitive processes alone but are also related to emotional responses and behavioral actions.

From a neural basis perspective, executive functions have been shown to be largely localized in the prefrontal cortex. This region is associated with executive functions such as working memory, inhibitory control, and flexible shifting of attention (Fassbender et al., 2004; Scheibel et al., 2003). The prefrontal cortex regulates perception, thought, and behavior by activating and inhibiting other brain regions (Shallice, 2002). Executive functions developing in this region show significant progress particularly during the preschool period (Isquith et al., 2018; Sonuga-Barke et al., 2002). Research indicates that the core components of executive functions become distinguishable around the age of five (Garon et al., 2008).

- **Working Memory:** The capacity to temporarily hold, maintain, and manipulate information. It supports reasoning, language comprehension, and coordination across cognitive tasks, facilitates switching between multiple tasks, and prevents distractibility (Gathmann et al., 2015). In daily life, working memory plays a critical role in planning, sequencing tasks, and achieving individual goals (Alloway & Alloway, 2010; Kroesbergen et al., 2014).

- **Inhibitory Control (Inhibition):** The ability to suppress inappropriate automatic responses and focus attention on the primary goal (Diamond, 2013). This ability plays an important role in the regulation of behavior, emotion, and attention, and supports the coordination of executive functions by helping working memory filter out irrelevant information (Flook et al., 2015).

- **Cognitive Flexibility:** The ability to shift thoughts and perspectives, transition between different viewpoints, and adapt to changing conditions (Diamond, 2002). It is the latest-developing component among executive functions, and the ability to perform appropriately in complex tasks is associated with this process (Brooks et al., 2003).

3. Theoretical Foundations of Science Education in Early Childhood

The process of learning in science begins long before children enter formal education, and when appropriate learning environments are provided, individuals in early childhood can acquire fundamental scientific competencies (Koerber et al., 2005). Children naturally enjoy exploring scientific phenomena, and these early gains provide significant advantages in their later academic lives (Oppermann et al., 2018; Sackes et al., 2011). This situation has led to a strong consensus in the academic literature regarding the necessity of addressing science education as an integral part of early childhood curricula (Brenneman et al., 2009).

Early science education is regarded as a means of fostering children's natural curiosity, motivation, and problem-solving skills (Clements & Sarama, 2016). Children display an instinctive tendency to explore their environment and demonstrate high levels of motivation during this process (Raffini, 1993). In this context, science activities in the preschool period shape children's early experiences with scientific phenomena and establish the cognitive and motivational foundations that enable them to develop a "science learner" identity (Aschbacher et al., 2010).

The importance of science education at an early age necessitates an integrated consideration of conceptual and procedural knowledge dimensions. While conceptual knowledge enables children to understand environmental phenomena, procedural knowledge supports problem-solving abilities by developing "doing science" skills such as observation, prediction, and inference (Eshach & Fried, 2005). The development of positive attitudes toward science education among children depends on teachers' ability to present science topics in ways that stimulate curiosity and increase engagement. To this end, teachers should support science activities with real-life examples and hands-on experiences, providing interactive and enjoyable learning environments that promote learning by doing and experiencing (Oppermann et al., 2018).

Historically, approaches to preschool science education have been shaped

by Piaget’s constructivist theory of cognitive development and Vygotsky’s emphasis on social interaction (National Association for the Education of Young Children [NAEYC], 2009). This perspective defines one of the primary goals of preschool education as ensuring children’s cognitive and social readiness for primary school. However, there are differences of opinion among educators regarding the definition of “readiness” and the means of achieving it; while some advocate for an academically focused approach, others adopt a whole-child perspective and propose a holistic educational framework that supports not only cognitive development but also physical and social-emotional development (Bishop-Josef & Zigler, 2011; Wilinski, 2017).

Preschool science education aims not only to promote cognitive development but also to enhance children’s scientific process skills, problem-solving abilities, and capacity to address everyday problems using scientific methods. Nevertheless, the time allocated to science content in existing curricula is limited, and teachers’ insufficient pedagogical and content knowledge in science hinders the provision of adequate emphasis on science education (Oppermann et al., 2018; Piasta et al., 2014).

In recent years, paradigms in science education have adopted a holistic approach that encompasses not only individual cognitive development but also social, cultural, and motivational dimensions. This reform-oriented perspective encourages learners not merely to acquire knowledge but to understand and apply science (Morris, 2025). The preschool period is considered a critical stage for the implementation of this holistic approach, as it supports children’s curiosity and motivation while ensuring the development of scientific process skills, problem-solving abilities, and creative thinking capacities.

4. The Interaction Between Science Education and Executive Functions

The theoretical foundations of science education in early childhood are addressed within the framework of children’s cognitive developmental characteristics and, in particular, the role of executive function capacity in scientific thinking processes. It is emphasized that executive function capacity plays a critical role in preschool children’s processes of testing hypotheses and evaluating obtained results, and that children are able to test hypotheses even before they fully possess the capacity to confirm or revise them (Gropen et al., 2011). This highlights that science education in early childhood should be structured not merely as a process of knowledge transmission, but also as a learning domain that supports cognitive control mechanisms.

The development of executive function capacity is age-related and largely based on the maturation of the prefrontal cortex (Zelazo et al., 2003; Zelazo et al., 2008). However, executive function capacities in preschool children are not solely the result of biological maturation; through appropriately

structured pedagogical approaches, these capacities can be actively integrated into learning processes (Madanipour et al., 2025). Research demonstrates that executive function skills in the preschool period can be strengthened through educational programs and structured activities, indicating that children require pedagogical support in order to use their cognitive control processes effectively within learning contexts (Blair, 2016). In particular, when children are unfamiliar with specific science content, they may be unable to use inhibitory control effectively due to limited attentional capacity and working memory resources, and may struggle to coordinate these processes efficiently (Espy & Bull, 2005). In this regard, early childhood science education should be designed in ways that support executive function components such as attention, working memory, and inhibitory control.

Studies in the literature indicate that executive function capacity can be enhanced through education (Diamond et al., 2007; Dowsett & Livesey, 2000; Kloo & Perner, 2003; Landry et al., 2002; Thorell et al., 2009). These findings challenge a common assumption regarding early science education: The view that it is meaningless to attempt to change preschool children's scientific misconceptions. On the contrary, the malleability of executive functions suggests that children can gradually become individuals with higher levels of cognitive control and develop stronger critical thinking skills (Klaczynski & Cottrell, 2004; Stanovich, 1999; Stanovich & West, 2000). In this context, scientific thinking involves not only possessing accurate scientific concepts but also developing mental habits that support cognitive control within scientific inquiry processes.

This developmental process related to science education can be understood more clearly within the framework of levels of Reflective Consciousness (RC). According to Zelazo and colleagues (2006), children between the ages of 3 and 4 are typically at the level of Reflective Consciousness 1 (RC1), at which they are able to use single-dimensional rule pairs. That is, they can apply rules that make distinctions within the same dimension. For example, within the Dimensional Change Card Sort (DCCS) task, three-year-old children can sort two-dimensional picture sets according to only one dimension (e.g., color). However, they are unable to shift between dimensions (Jacques et al., 1999; Zelazo, 2006). This indicates that children can think about one rule in relation to another but cannot integrate incompatible rule pairs within a single system. Only when children reach Reflective Consciousness 2 (RC2) at around four years of age are they able to revise and verify their predictions through hypothesis testing. At this level, the ability to integrate incompatible rule pairs within a single system develops, and children are not only able to make predictions but can also explain why and how they changed their predictions based on the results obtained (Zelazo et al., 2006). This developmental distinction clearly demonstrates why pedagogical approaches

that are sensitive to executive function capacity are critical during science education.

Within this theoretical framework, the cycle of hypothesis testing and revision constitutes a critical starting point for preschool science education. This cycle enables teachers to make children's existing knowledge and thinking visible and to use this information to deepen children's scientific understanding. Preschool science learning can occur at different levels of cognitive complexity that involve the application and strengthening of executive function capacity. At the RC1 level, teachers can support children's use of executive functions by encouraging focused predictions and accurate observations in a new science domain, while also enriching the process through reflective activities. At the RC2 level, the goal is not only for children to abandon incorrect conceptions but also to be able to explain why and how they revised their predictions. This process supports the development of language and metacognitive awareness in addition to cognitive control (Gropen et al., 2011).

From the perspective of the theoretical foundations of early childhood science education, reflective processes occupy an important place. Comparing discrepancies between children's predictions and observations, emphasizing the dimensions associated with accurate predictions, and creating collaborative learning environments are among effective classroom practices that support reflection (Gopnik & Astington, 1988; Gropen et al., 2011; Kloos & Van Orden, 2005). Such practices allow children to deepen their scientific sense-making processes by comparing their own thinking with that of others.

Finally, the theoretical foundations of science education in early childhood are also directly related to the level of science content knowledge that teachers possess. Teachers' use of content knowledge varies depending on students' executive function capacities. At the RC1 level, content knowledge is indispensable for planning investigations that support specific predictions, require accurate observations, and challenge children's misconceptions. At the RC2 level, additional content knowledge enables teachers to engage in longer and more in-depth interactions with children about scientific phenomena and to use appropriate scientific language. In this context, language emerges as a fundamental tool for making children's implicit thinking explicit (Karmiloff-Smith, 1996; Zelazo et al., 2006). When appropriately structured, explanation and reflection are regarded as powerful pedagogical tools that support children in articulating their thoughts, forming new connections, and deepening scientific knowledge (Bowman et al., 2001; Epstein, 2003; Landry & Forman, 1999).

5. Conclusion and Recommendations

The interaction between science education and executive functions in early childhood plays a central role in shaping children's cognitive, social,

and motivational development. Executive functions, particularly components such as working memory, inhibitory control, and cognitive flexibility, serve fundamental roles in children's problem-solving, hypothesis testing, and scientific thinking processes (Diamond, 2013; Zelazo & Carlson, 2012). These functions not only develop through biological maturation but can also be strengthened through appropriate pedagogical interventions and strategies, and effectively integrated into children's learning processes.

In this context, it is crucial that the pedagogical strategies of early childhood science education be designed in accordance with children's levels of cognitive development. Children between the ages of 3 and 4 are generally at the RC1 level, where they can make single-dimensional predictions and organize their observations using limited rules. During this period, it is recommended that teachers design activities that direct children's attention toward a specific goal and encourage them to use their executive functions through simple experiments. For instance, children's observational skills can be enhanced through experimental questions such as whether colored objects sink or float in water.

Older children, upon reaching the RC2 level, develop the capacity to integrate conflicting rules and to revise their predictions. At this stage, teachers should organize activities that guide children not only to correct misconceptions but also to explain these concepts within a cause-and-effect framework. Such activities promote children's cognitive control and metacognitive awareness while also facilitating the acquisition of higher-level scientific thinking skills.

The pedagogical implementation of science education should support hands-on learning and make learning experiences interactive and inquiry-based. Teachers' use of science content knowledge in a manner that is sensitive to children's executive function capacities is a determining factor in creating effective learning environments. Moreover, the systematic integration of explanation and reflection processes will enable children to compare their own thinking with that of others and to deepen their scientific sense-making processes.

In conclusion, early childhood science education not only provides scientific knowledge but also offers a holistic learning experience that supports the development of executive functions, particularly key components such as cognitive control, problem-solving, and metacognitive awareness. In this process, teachers should structure activities considering children's cognitive capacities, use content knowledge as a guiding tool, and systematically integrate reflective processes. In this way, children can develop as individuals competent in both scientific knowledge and cognitive control processes.

References

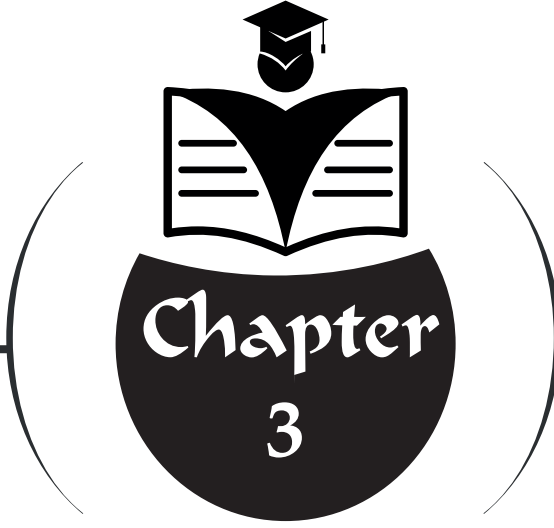
- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106*(1), 20-29.
- Anderson, P. (2002). Assessment and development of Executive Function (EF) during childhood. *Child Neuropsychology, 8*(2), 71-82.
- Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching, 47*(5), 564-582.
- Bishop-Josef, S. J., & Zigler, E. (2011). The cognitive/academic emphasis versus the whole child: The formative years. In E. Zigler, W. S. Gilliam, & S. M. Jones (Eds.), *A vision for universal preschool education* (pp. 67-84). Cambridge, MA: Cambridge University Press.
- Blair, C. (2016). Executive function and early childhood education. *Current Opinion in Behavioral Sciences, 10*, 102-107.
- Blair, C. (2017). Educating executive function. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(1-2), e1403.
- Bowman, B., Donovan, M. S., & Burns, M. S. (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.
- Brenneman, K., Stevenson-Boyd, J., & Frede, E. C. (2009). Math and science in preschool: Policies and practice. *Preschool Policy Brief, 19*, 1-12.
- Brooks, P. J., Hanauer, J. B., Padowska, B., & Rosman, H. (2003). The role of selective attention in preschoolers' rule use in a novel dimensional card sort. *Cognitive Development, 18*(2), 195-215.
- Clements, J., & Sarama, J. (2016). *Learning and teaching early math: The learning trajectories approach*. New York, NY: Routledge.
- Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology, 10*(3), e1001293.
- National Association for the Education of Young Children (NAEYC) (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. Retrieved January 20, 2026, from <https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/PSDAP.pdf>
- Corcoran, R. P., & O'Flaherty, J. (2017). Executive function during teacher preparation. *Teaching and Teacher Education, 63*, 168-175.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466-503). New York, NY: Oxford University Press.

- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64(1), 135-168.
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, 318(5855), 1387-1388.
- Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of “executive skills” training. *Developmental Psychology: The Journal of the International Society for Developmental Psychobiology*, 36(2), 161-174.
- Elliott, R. (2003). Executive functions and their disorders. *British Medical Bulletin*, 65(1), 49-59.
- Epstein, A. S. (2003). How planning and reflection develop young children’s thinking skills. *Young Children*, 58(5), 28-36.
- Eshach, H., & Fried, M. N. (2005). Should science be taught in early childhood? *Journal of Science Education and Technology*, 14(3), 315-336.
- Espy, K. A., & Bull, R. (2005). Inhibitory processes in young children and individual variation in short-term memory. *Developmental Neuropsychology*, 28(2), 669-688.
- Fassbender, C., Murphy, K., Foxe, J. J., Wylie, G. R., Javitt, D. C., Robertson, I. H., & Garavan, H. (2004). A topography of executive functions and their interactions revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, 20(2), 132-143.
- Flook, L., Goldberg, S. B., Pinger, L., & Davidson, R. J. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Developmental Psychology*, 51(1), 44-51.
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134(1), 31-60.
- Gartstein, M. A., Bridgett, D. J., Young, B. N., Panksepp, J., & Power, T. (2013). Origins of effortful control: Infant and parent contributions. *Infancy*, 18(2), 149-183.
- Gathmann, B., Schiebener, J., Wolf, O. T., & Brand, M. (2015). Monitoring supports performance in a dual-task paradigm involving a risky decision-making task and a working memory task. *Frontiers in Psychology*, 6, 142.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test review behavior rating inventory of executive function. *Child Neuropsychology*, 6(3), 235-238.
- Gopnik, A., & Astington, J. W. (1988). Children’s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59(1), 26-37.
- Gropen, J., Clark-Chiarelli, N., Hoisington, C., & Ehrlich, S. B. (2011). The importance of executive function in early science education. *Child Development Perspectives*, 5(4), 298-304.
- Groppe, K., & Elsner, B. (2015). The influence of hot and cool executive function on

- the development of eating styles related to overweight in children. *Appetite*, 87, 127-136.
- Hendry, A., Jones, E. J., & Charman, T. (2016). Executive function in the first three years of life: Precursors, predictors and patterns. *Developmental Review*, 42, 1-33.
- Isquith, P. K., Gioia, G. A., & Espy, K. A. (2018). Executive function in preschool children: Examination through everyday behavior. In *Using developmental, cognitive, and neuroscience approaches to understand executive control in young children* (pp. 403-422). New York, NY: Psychology Press.
- Jacobson, L. A., Koriakin, T., Lipkin, P., Boada, R., Frijters, J. C., Lovett, M. W., ... & Mahone, E. M. (2017). Executive functions contribute uniquely to reading competence in minority youth. *Journal of Learning Disabilities*, 50(4), 422-433.
- Jacques, S., Zelazo, P. D., Kirkham, N. Z., & Semcesen, T. K. (1999). Rule selection versus rule execution in preschoolers: An error-detection approach. *Developmental Psychology*, 35(3), 770-780.
- Karmiloff-Smith, A. (1996). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Klaczynski, P. A., & Cottrell, J. M. (2004). A dual-process approach to cognitive development: The case of children's understanding of sunk cost decisions. *Thinking & Reasoning*, 10(2), 147-174.
- Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development*, 74(6), 1823-1839.
- Kloos, H., & Van Orden, G. C. (2005). Can a preschooler's mistaken belief benefit learning? *Swiss Journal of Psychology*, 64(3), 195-205.
- Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: Preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology*, 64(3), 141-152.
- Korkmaz, B., Njiokiktjien, C., & Verschoor, C. A. (2013). *Children's social relatedness: An embodied brain process*. Amsterdam, Netherlands: Suyi Publications.
- Kroesbergen, E. H., van 't Noordende, J. E., & Kolkman, M. E. (2014). Training working memory in kindergarten children: Effects on working memory and early numeracy. *Child Neuropsychology*, 20(1), 23-37.
- Landry, C. E., & Forman, G. E. (1999). Research on early science education. In C. Seefeldt (Ed.), *The early childhood curriculum: Current findings in theory and practice* (pp. 133-158). New York, NY: Teachers College Press.
- Landry, S. H., Miller-Loncar, C. L., Smith, K. E., & Swank, P. R. (2002). The role of early parenting in children's development of executive processes. *Developmental Neuropsychology*, 21(1), 15-41.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59-80.

- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.
- Lunt, L., Bramham, J., Morris, R. G., Bullock, P. R., Selway, R. P., Xenitidis, K., & David, A. S. (2012). Prefrontal cortex dysfunction and 'Jumping to Conclusions': Bias or deficit? *Journal of Neuropsychology*, 6(1), 65-78.
- Madanipour, P., Garvis, S., Cohrssen, C., & Pendergast, D. (2025). Early childhood teachers' understanding of executive functions and strategies employed to facilitate them. *Sec. Teacher Education*, 9.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
- Morris, D. L. (2025). Rethinking science education practices: Shifting from investigation-centric to comprehensive inquiry-based instruction. *Education Sciences*, 15(1), 73.
- Oppermann, E., Brunner, M., Eccles, J. S., & Anders, Y. (2018). Uncovering young children's motivational beliefs about learning science. *Journal of Research in Science Teaching*, 55(3), 399-421.
- Piasta, S. B., Logan, J. A. R., Pelatti, C. Y., Capps, J. L., & Petrill, S. A. (2015). Professional development for early childhood educators: Efforts to improve math and science learning opportunities in early childhood classrooms. *Journal of Educational Psychology*, 107(2), 407-422.
- Raffini, J. P. (1993). *Winners without losers: Structures and strategies for increasing student motivation to learn*. Boston, MA: Allyn & Bacon.
- Sackes, M. (2013). Children's competencies in process skills in kindergarten and their impact on academic achievement in third grade. *Early Education & Development*, 24(5), 704-720.
- Salimpour, V. N., & Desrocher, M. (2006). Increasing the Utility of EF Assessment of Executive Function in Children. *Developmental Disabilities Bulletin*, 34, 15-42.
- Scheibel, R. S., Pearson, D. A., Faria, L. P., Kotrla, K. J., Aylward, E., Bachevalier, J., & Levin, H. S. (2003). An fMRI study of executive functioning after severe diffuse TBI. *Brain Injury*, 17(11), 919-930.
- Shallice, T. (2002). Fractionation of the supervisory system. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 261-277). New York, NY: Oxford University Press.
- Sonuga-Barke, E. J. S., Dalen, L., Daley, D., & Remington, B. (2002). Are planning, working memory, and inhibition associated with individual differences in preschool ADHD symptoms? *Developmental Neuropsychology*, 21(3), 255-272.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.

- Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, 23(5), 701-717.
- Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 12(1), 106-113.
- Wilinski, B. (2017). *When pre-K comes to school: Policy, partnerships, and the early childhood education workforce*. New York, NY: Teachers College Press.
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297-301.
- Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, 6(4), 354-360.
- Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (2nd ed., pp. 553-574). Cambridge, MA: MIT Press.
- Zelazo, P. D., Gao, H. H., & Todd, R. (2006). The development of consciousness. In P. D. Zelazo, M. M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 403-430). Cambridge, MA: Cambridge University Press.
- Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, 68(3), vii-137.



ARTIFICIAL INTELLIGENCE BASED SOLUTION STRATEGIES IN SCIENCE EDUCATION: ANALYSIS OF HSTS QUESTIONS WITH CHATGPT*



Buğra Taşkıran¹

Mustafa Ergun²

1 M. Sc., Ondokuz Mayıs University, Graduate School of Education, Department of Science Education <https://orcid.org/0009-0002-7323-8290>

2 Assoc. Prof. Dr., Ondokuz Mayıs University, Faculty of Education, Department of Science Education, Samsun, Türkiye <https://orcid.org/0000-0003-4471-6601>

* This study was produced from the first author's master's thesis under the supervision of the second author.

Artificial Intelligence Performance in Exams

Artificial intelligence (AI) systems' performance in national examinations has garnered significant attention in recent years, particularly in fields such as healthcare and education. Meta analytical studies examining ChatGPT's performance in national health licensing examinations have indicated that version updates contribute to increased accuracy rates (Jin et al., 2024). Furthermore, within the context of health licensing examinations, ChatGPT models have been reported to demonstrate greater proficiency in the fields of pharmacy, medicine, dentistry, and nursing, respectively (Jin et al., 2024).

In the field of medicine, the performance of different ChatGPT models (3.5, 4, 4o) has been evaluated in examinations conducted across various countries (Alfershofer et al., 2023), including China (Ming et al., 2024; Zong et al., 2023), the United States (Gilson et al., 2023), Taiwan (Hsieh et al., 2024; Huang et al., 2024), India (Gandhi et al., 2024), Poland (Siebielec et al., 2024), Japan (Sawamura et al., 2024), United Kingdom (Giannos & Delarus, 2023) Sweden (Arvidsson et al., 2024), Spain (Solmonovich et al., 2024), France (Guigue et al., 2024), Italy (Rosettini et al., 2024), Türkiye (Oztermeli & Oztermeli, 2023), Brazil (Alessi et al., 2024) and Peru (Flores-Cohaila et al., 2023).

Although ChatGPT performs quite successfully on simple tasks at the basic knowledge and comprehension level, it has been determined that its performance declines significantly as the questions become more difficult. In particular, it has been found to be insufficient in solving complex problems that require application and in interpreting visual elements (such as graphs and tables) (Dao & Le, 2023).

GPT-4 demonstrates a high level of proficiency in providing conceptual explanations in physics. However, it remains comparatively weak in performing formal derivations and handling multi step symbolic manipulations (Bubeck et al., 2023). It has been determined that the mathematical problem solving performance of large language models such as GPT-4 varies significantly depending on the type and difficulty level of the problem. While these models tend to perform well on questions that require recalling definitions and applying fundamental knowledge, they have been found to be markedly insufficient when dealing with problems that require formal proofs and advanced reasoning (Frieder et al., 2023).

Meta analyses in the literature indicate that ChatGPT significantly enhances students' academic performance, while having moderately positive effects on their learning perceptions and higher-order thinking skills (Wang & Fan, 2025). However, this effect has been shown to vary depending on how ChatGPT is used, the type of course, and the duration of use. It was found that students who used ChatGPT in a teacher education program achieved higher

scores both in higher-order thinking skills and overall performance (Lu et al., 2024).

In standardized examinations like SAT, the performance of artificial intelligence varies considerably depending on the model generation, the difficulty level of the questions, and the subject domain. GPT-4 achieved an accuracy rate of 95% on the SAT Mathematics exam and 92.5% on the Gaokao English exam, thereby surpassing the average human performance on the SAT and LSAT (Zhong et al., 2024).

ChatGPT's SAT math performance has been reported to be moderate and variable. The model's overall accuracy rate in math is around 56%, ranging from 72% in some test sets to as low as 31% in others (Tan et al., 2024). It has been observed that the model performs better on multiple choice questions, but makes more errors on open-ended questions. Furthermore it has been determined that the model's performance is sensitive to the prompts used. It has been found that while instructions such as "just give the answer" reduce accuracy, allowing the student to generate solutions yields more accurate results. Overall, while ChatGPT demonstrates average success in high school level mathematics, it shows limitations in terms of consistency and deep problem solving (Tan et al., 2024).

AI systems generally demonstrate high performance in nationally standardized examinations. Particularly in structured and well-defined problem domains, they achieve results comparable to or even exceeding human performance. However, their performance tends to vary in areas that require open-ended responses and complex problem solving skills.

High School Transition System

The High School Transition System (HSTS) is an examination system implemented by the Ministry of National Education (MoNE) of the Republic of Türkiye since 2018. Based on a centralized exam score, it is designed for 8th-grade middle school students to be placed into high schools. This system enables student placement through both local placement and admission to schools that accept students via centralized exam scores. Following the announcement of the central exam scores, student placements are carried out initially based on their central exam results, followed by the local placement system. The HSTS examination is administered in two separate sessions on the same day, one in the morning and the other at noon. The first session encompasses tests on verbal subjects, while the second session includes tests on quantitative subjects, specifically science and mathematics (MoNE, 2025).

Table 1. HSTS Science Curriculum Topic Classification

| Discipline | Learning Domain | Unit Titles | Topics Covered |
|------------|-------------------------------------|--|---|
| Physics | Physical Phenomena; Earth and Space | Seasons and Climate, Pressure, Simple Machines, Electricity | Formation of Seasons, Climate and Air Movements, Pressure, Simple Machines, Electric Charges and Electrification, Electrically Charged Objects, Transformation of Electrical Energy |
| Chemistry | Structure and Properties of Matter | Matter and Industry | Periodic Table, Physical and Chemical Changes, Chemical Reactions, Acids and Bases |
| Biology | Living Organisms and Life Processes | DNA and Genetic Code; Energy Transformations and Environmental Science | DNA and Genetic Code, Heredity, Mutation and Modification, Adaptation, Biotechnology, Food Chain and Energy Flow, Energy Transformations, Matter Cycles, Sustainable Development |

The science test is prepared based on the 8th-grade curriculum. The Turkish HSTS science curriculum is organized into three core scientific disciplines: physics, chemistry, and biology, each aligned with specific learning domains and unit structures (Table 1). Physics encompasses topics within the domains of physical phenomena and earth-space science, such as seasonal changes, pressure mechanics, simple machines, and electricity, all of which emphasize conceptual understanding and application in everyday contexts. Chemistry centers on the structure and properties of matter, focusing on units like the periodic table, chemical reactions, and acid-base behavior, thereby fostering foundational knowledge of matter transformations. Biology is grounded in studying living organisms and life processes, emphasizing genetics, biotechnology, ecology, and sustainability. The distribution reflects a curriculum design that balances disciplinary content with interdisciplinary connections, preparing students to engage in complex scientific reasoning and real world problem solving.

Table 2. HSTS Science Exam Topic-Based Question Distribution

| Unit Title | Average Number of Questions | Description |
|-------------------------------------|-----------------------------|--|
| Seasons and Climate | 1–2 questions | Predominantly includes questions requiring interpretation of graphs and data. |
| DNA and Genetic Code | 2–3 questions | Emphasizes topics such as heredity and genetic mutations. |
| Pressure | 2 questions | Typically integrates solid, liquid, and gas pressure concepts within a unified question structure. |
| Periodic Table – Chemical Reactions | 2–3 questions | Focuses on the structure and transformation of matter. |

| Unit Title | Average Number of Questions | Description |
|--|-----------------------------|--|
| Simple Machines | 1–2 questions | Assesses understanding of mechanical systems in real-life contexts. |
| Electric Charges and Electrical Energy | 2 questions | Frequently addresses circuit components and the interaction between electric charges. |
| Energy Transformations and Environmental Science | 2–3 questions | Covers core concepts such as food chains, energy flow, and sustainability. |
| Acids, Bases, and Physical-Chemical Changes | 1–2 questions | Assesses identification of acids and bases through practical, everyday-life scenarios. |

Table 2. reflects the approximate thematic weight of each unit in the HSTS science component, indicating the prioritized content areas and associated cognitive demands. The exam questions are designed to assess not only students' knowledge levels but also their higher order cognitive skills, such as reasoning, scientific process skills, interpretation of graphs and tables, and problem based thinking. The question format is multiple choice, each item offering four answer options. These questions typically require reading comprehension and interpretative skills. They often incorporate experimental scenarios and contexts related to daily life. Additionally, skill based questions may include visual elements, graphs, and tables. There are also questions that consist solely of textual content.

When designing HSTS science questions, the Turkish MoNE adopts an assessment approach in line with international frameworks such as PISA and TIMSS. This orientation aims not merely to evaluate students' factual knowledge but to assess their ability to apply scientific understanding to problem-solving contexts.

Table 3. HSTS Science Exam Classification by Question Type

| Question Type | Description | Approximate Proportion (out of 20 questions) |
|--|--|--|
| Interpretation, Graph and Table Analysis | Requires analysis and inference based on provided graphs, tables, or visual data. Frequently utilized in units such as environmental science, DNA, and pressure. | 30–35% (6–7 questions) |
| Scientific Process Skills (Experiment, Hypothesis) | Involves scientific reasoning such as forming hypotheses, identifying variables, and interpreting experimental scenarios. | 20–25% (4–5 questions) |
| Contextual Questions Based on Real-Life Scenarios | Assesses the application of scientific knowledge within everyday contexts. Common in topics such as electricity, energy conversions, and acid-base chemistry. | 15–20% (3–4 questions) |

| Question Type | Description | Approximate Proportion (out of 20 questions) |
|---|---|--|
| Conceptual Knowledge, Definition-Based Items | Relies on direct recall of factual knowledge, concept identification, or error detection. Less frequently emphasized due to the exam's focus on application of knowledge. | 10–15% (2–3 questions) |
| Causal Reasoning, Logical Inference Questions | Requires establishing cause-effect relationships and analyzing processes within a logical framework. Often used in biology and physical sciences. | 15–20% (3–4 questions) |

This table 3 reflects the cognitive and analytical depth of the Turkish national HSTS science assessment, emphasizing scientific literacy, data interpretation, and applied reasoning over rote memorization. A significant proportion of the questions are constructed to require multi step reasoning, establishing conceptual relationships, and analyzing distractors, thereby encouraging deeper cognitive engagement. Furthermore, the prevalence of simple recall based questions is relatively low, necessitating students to demonstrate a higher degree of scientific literacy, critical thinking, and analytical skills to succeed.

The aim of the study

This study aims to examine the performance of ChatGPT-4 in solving the High School Transition System science questions and how it explains the answers to these questions. Accordingly, ChatGPT-4's language processing capabilities, image processing success, knowledge level in science subjects, appropriateness to student level, and problem solving strategies were analyzed in detail.

Based on the identified research problem, the sub-problems are as follows:

1. What is the accuracy rate of the answers provided by ChatGPT-4 to HSTS science exam questions, and in which types of questions do incorrect responses most commonly occur?
2. To what extent is ChatGPT-4 capable of understanding HSTS science question texts and extracting the necessary information?
3. How successful is ChatGPT-4 in processing and interpreting visual elements included in HSTS science questions?
4. How effective and detailed are the explanations provided by ChatGPT-4 for its answers to HSTS science questions?
5. To what degree are ChatGPT-4's responses to LGS Science questions appropriate for the students' grade level?

The methodology of the study

In this study, the case study design was employed as a qualitative research method. A case study aims to examine one or a few cases in depth and evaluates the factors related to a case through a holistic perspective (Yıldırım & Şimşek, 2021). This method is preferred when investigating contemporary phenomena within their real-life context, particularly in situations where the boundaries between the phenomenon and its context are not clearly evident and where multiple sources of evidence are available (Yıldırım & Şimşek, 2021).

The case study approach is particularly suitable for thoroughly examining how ChatGPT-4 approaches HSTS science questions, its problem solving processes, and the challenges it encounters. It enables a comprehensive understanding of the solution strategies developed in response to the specific problems faced by the model and the effectiveness of these strategies (Yıldırım & Şimşek, 2021).

The content analysis method was used to analyze the data. Content analysis is suitable for evaluating the quality, scope, and accuracy of ChatGPT-4's answers to HSTS science questions, and it allows for a detailed examination of the model's language processing capacity and its mastery of the subject matter. For each question, the accuracy of ChatGPT-4's responses was checked and its solution processes were examined. The model's rate of correct answers was considered the primary indicator of its performance. Factors such as the types of questions it answered correctly, the complexity of the questions, the clarity of its explanations, and their appropriateness to the student level played a significant role in determining the model's overall success. Questions that were answered incorrectly were analyzed in detail to identify the types of questions the model struggled with and to examine the underlying reasons for these errors. This analysis provided valuable insights into the model's potential use as an instructional tool and areas for improvement in terms of teaching methodologies. The themes and codes derived from the content analysis are presented in Table 4, and the findings were interpreted accordingly.

Table 4. Themes and Codes

| Themes | Codes |
|-----------------------------------|--|
| Accuracy and Incorrectness | Correct answer Incorrect answer |
| Types of Incorrect Answers | Misunderstanding the question Failure to relate information Visual content related Text based |
| Visual Processing | Successful processing Unsuccessful processing |

| Themes | Codes |
|---|-----------------------------------|
| Effectiveness of Explanation | Highly effective explanations |
| | Moderately effective explanations |
| | Ineffective explanations |
| Appropriateness to Student Level | Appropriate |
| | Not appropriate |

A total of 120 questions from the science sections of the HSTS exams administered between 2018 and 2023 were identified and categorized. The questions were presented to the AI model using different methods depending on the type of question. Screenshots of the visuals, graphs, and tables included in the questions were taken and uploaded to the ChatGPT-4. The textual parts of the questions were transcribed verbatim and submitted to the model along with the visuals. For each question, prompts appropriate to its content were prepared, and ChatGPT-4 was asked to solve them. The prompts were designed in a way that students might naturally write them. Sample prompts are provided in Table 5.

Table 5. Sample Prompts Written to the ChatGPT

| Sample Prompts Written to the ChatGPT |
|---|
| The diagram in the question shows identical and fraternal twins. Accordingly, what is the correct answer to the question? |
| The image includes a crossbreeding experiment consisting of two stages. Can you solve the question based on this? |
| The first image shows the experimental materials mentioned in the question. The second image presents the answer choices consisting of experimental setups made from these materials. Accordingly, which option is correct? |
| Can you solve the question for me by explaining all the answer choices? |
| What is the correct answer to the question? |

Findings

Each HSTS exam contains 20 questions in the field of science. Some of these questions include visual content (tables, graphs, and figures), while others are text-based questions only. The distribution of questions containing visual content (tables, graphs, and figures) and text-based questions only, by year, is shown in Figure 1.

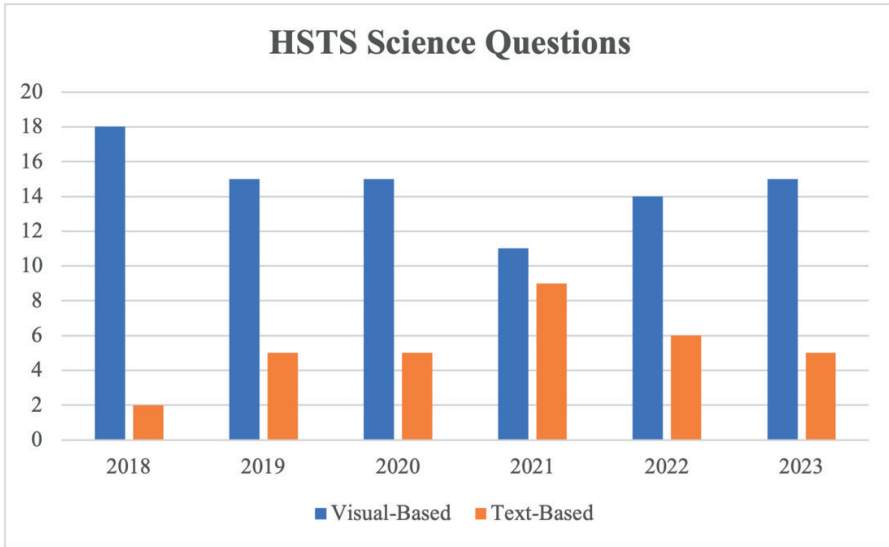


Figure 1. *Distribution of visual and text-based questions by year*

Between 2018 and 2023, 67% of the total 120 HSTS science questions, or 88 questions, were visual-based, while 33%, or 32 questions, were text-based. Visual-based questions consistently outweighed text-based questions in all years. In 2018, there were 18 questions with visual content, while there were only 2 text-based questions. In 2019 and 2020, there were 15 questions with visual content and 5 text-based questions. In 2021, the number of visual questions dropped to 11, while text-based questions rose to 9, almost evenly distributed. In 2022, visual questions were recorded as 14 and text-based questions as 6. In 2023, visual questions were 15 and text-based questions were 5.

These data indicate that visual questions are generally dominant, while text-based questions remain fewer but at a relatively stable rate. It is understood that in the 2021 exam, the number of both question types became almost equal and a change was made in the exam design. Apart from this exception, visually based questions are predominantly preferred in the national exam.

ChatGPT's Effectiveness in Visual Based Questions

Figure 2 shows the distribution of correct, incorrect, and unanswered responses provided by ChatGPT over the years to science questions that only included visual content (tables, graphs, and diagrams).

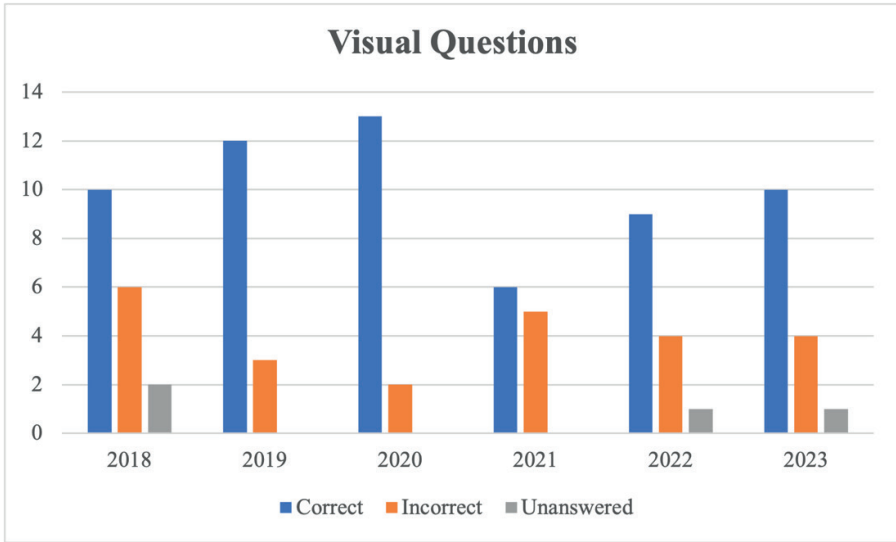



Figure 2. *Distribution of correct and incorrect answers to visual based questions by year*

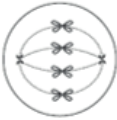
In 2018, out of a total of 18 questions with visual content, 10 correct answers, 6 incorrect answers, and 2 unanswered questions were recorded. In 2019, correct answers increased to 12, incorrect answers decreased to 3, and there were no unanswered questions. In 2020, correct answers reached their highest level at 13, incorrect answers were 2, and there were no unanswered questions. In 2021, correct answers dropped to 9, incorrect answers were recorded as 6, and there were no unanswered questions. In 2022, correct answers rose to 11, incorrect answers were 4, and unanswered questions were 1. In 2023, correct answers were 10, incorrect answers were 4, and unanswered questions were 1. ChatGPT answered 61 out of 88 visually based questions correctly, provided incorrect answers to 23, and left only 4 unanswered. These results indicate an accuracy rate of 69%, an error rate of 26%, and an unanswered rate of 5% for visually based questions.


2. Bir hayvanda;


- yaraların iyileşmesi,
- embriyonun gelişmesi

olayların gerçekleşmesi sırasında aşağıdaki hücre bölünmesi evrelerinden hangisi görülmez?

A) 

B) 

C) 

D) 

ChatGPT

Something went wrong. If this issue persists please contact us through our help center at help.openai.com.

There was an error generating a response

Regenerate

ChatGPT can make mistakes. Consider checking important information.


Figure 3. Question 2 for 2018 and ChatGPT's answer

In the question in Figure 3, it is asked which of the cell division stages observed during events such as wound healing and embryo development in an animal will not be seen. The expected answer is to select the correct mitosis stage and eliminate the incorrect image. However, ChatGPT made a mistake in this visual-based question. To determine whether this was a one time error or a recurring issue, a new conversation was initiated and the question was asked again, resulting in the same error message.

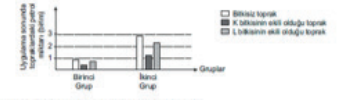
6. Birim insanları, tozuma sırasında dökülen petrolün toprakta oluştuğu kirliliğin K ve L bitkileri kullanarak azaltılabileceğini göstermek amacıyla bir proje hazırlıyorlar. Birim insanları, dökülen petrolü bitkiler kullanarak ortamdaki uzaklaştırma kapasitesini bu bitkilerin genlerine daha hızlı büyüyen bitkilere aktaracaklar. Elde edilen genetik değişimleri bu bitkilerde petrolü toprakları daha hızlı bir şekilde uzaklaştırmak için kullanacaklar.

Bu proje kapsamında aşağıdaki işlemler gerçekleştiriliyor:

- Aile üseti özdey toprak alan seçilir bunlardan iki grup oluşturuluyor.
- Popülasyon birinci gruptaki üç özdey toprak alanına birer bitim, ikinci gruptaki üç özdey toprak alanına da üçer bitim konuluyor.



Uygulama sonunda, topraklarda kalan bu petrolün miktarları grafikteki gibidir:



Buna göre aşağıdaki ifadelerden hangisi söylenemez?

A) K ve L bitkileri, petrolün ortamdaki uzaklaştırılmasında hiç etkili olmadığı için daha hızlı büyüyen bitkiler seçilmelidir.

B) K bitkisinin petrolün ortamdaki uzaklaştırılmasında sorumlu olan genlerinin hızlı büyüyen diğer bitkilere aktarılması daha uygundur.

C) L bitkisinin petrolün ortamdaki uzaklaştırılmasında sorumlu olan genlerinin hızlı büyüyen diğer bitkilere aktarılması daha uygundur.

D) L bitkisi çok hızlı büyüdüğü için petrolün ortamdaki uzaklaştırılmasında K bitkisininin daha etkili olmuştur.

ChatGPT

Grafikte, her iki grupta da K bitkisinin eklendiği topraklarda petrol miktarının önemli ölçüde azaldığı görülmektedir. Bu, K bitkisinin petrolü toprakları uzaklaştırmada etkili olduğunu göstermektedir. L bitkisinin eklendiğinde petrol miktarında herhangi bir önemli azalma görülmemiştir. Grafikte ayrıca, ikinci grupta (daha fazla petrol eklenen grupta) K bitkisi ile işlem gören topraklarda, birinci gruptaki benzer topraklara göre daha az petrol kaldığı görülmektedir; bu da K bitkisinin farklı konsantrasyonlardaki petrolü azaltmada etkili olduğunu göstermektedir.

Buna göre aşağıdaki ifadelerden hangisi doğru söylenebilir:

B) K bitkisinin petrolün ortamdaki uzaklaştırılmasında sorumlu olan genlerinin hızlı büyüyen diğer bitkilere aktarılması daha uygundur.

Bu ifade, K bitkisinin petrolü toprakları uzaklaştırma potansiyeline sahip olduğunu ve bu özelliğin, petrolü daha hızlı bir şekilde uzaklaştırmak için kullanılacak hızlı büyüyen diğer bitkilere aktarılmasının uygun olacağını öne sürer, ki bu da grafikte gösterilen verilere uygundur. Diğer seçenekler, elde edilen verilere dayanarak yanlış çıkarımlar içermektedir.

Figure 4. Question 6 for 2018 and ChatGPT's answer

Figure 4 shows a question asked in 2018 where ChatGPT was expected to solve the problem using both its graph reading and visual processing capabilities. This question presented two different experiment visuals and their results in graph form. The ability to correctly read these graphs and draw meaningful conclusions was used to evaluate ChatGPT's performance. ChatGPT understood the situations in the images and correctly interpreted the given graph, accurately evaluating the independent, dependent, and constant variables in the science experiments. Accordingly, it identified the correct answer and provided an explanation.

3. Fotosentezin yapay ışıkta gerçekleşip gerçekleşmediğini araştırmak isteyen bir öğrenci verilen malzemelerden uygun olanları seçerek bir deney düzenineği oluşturacaktır.

DENEY MALZEMELERİ

Öğrenci, güneş ışığı alan bir ortamda araştırma amacına yönelik tek bir deney düzenineği hazırlayarak düzenekteki oksijen miktarı değişimini gözlemleyebilir.

Bu öğrencinin araştırma amacına uygun olarak hazırladığı deney düzenineği aşağıdakilerden hangisi gibi olmalıdır? (Işık geçirmez kutular, içerisindeki düzeneklerin görülebilmesi için ön yüzleri opak gösterilebilir.)

A)

Işık geçirmez kutu

B)

Işık geçirmez kutu

C)

Şeffaf kutu

D)

Şeffaf kutu

ChatGPT 4

ChatGPT

Verilen görseller, fotosentezin yapay ışıkta gerçekleşip gerçekleşmediğini araştırmak isteyen bir öğrencinin kullanabileceği deney düzenekleri ve malzemeleri göstermektedir.

Deney malzemeleri şunları içerir:

- * Işık ve hava geçirmez kutu
- * Hava geçirmez şeffaf kutu
- * KOH çözeltisi (Karbon dioksit tutucudur.)
- * Beyaz ışık veren el feneri
- * K bitkisi
- * Ortamdaki O2 miktarını ölçen cihaz

Öğrencinin araştırma amacına yönelik tek bir deney düzenineği hazırlayarak düzenekteki oksijen miktarı değişimini gözlemlemesi gerekmektedir. Deney düzenekleri şunlardır:

A) Işık geçirmez kutu içerisinde K bitkisi, beyaz ışık veren el feneri ve O2 miktarını ölçen cihaz bulunmaktadır.

B) Işık geçirmez kutu içerisinde K bitkisi, KOH çözeltisi ve O2 miktarını ölçen cihaz bulunmaktadır.

C) Şeffaf kutu içerisinde K bitkisi ve O2 miktarını ölçen cihaz bulunmaktadır.

D) Şeffaf kutu içerisinde K bitkisi, beyaz ışık veren el feneri, KOH çözeltisi ve O2 miktarını ölçen cihaz bulunmaktadır.

Eğer öğrenci güneş ışığı alan bir ortamda deney yapacaksa, şeffaf kutuyu kullanarak güneş ışığının bitkiye ulaşmasını sağlaması gerekecektir. Ancak, yapay ışıkta fotosentezin gerçekleşip gerçekleşmediğini araştırmak istiyorsa, güneş ışığını dışarıda tutacak ışık geçirmez bir kutu ve yapay ışık kaynağı olarak beyaz ışık veren el fenerini kullanması gerekir. Ayrıca, fotosentez esnasında ortaya çıkan oksijen miktarını ölçmek için O2 ölçen cihazın da düzenineğin içerisinde olması gerekir.

Bu bilgiler ışığında, doğru düzenek KOH çözeltisi olmadan ışık geçirmez kutu, beyaz ışık veren el feneri v. ↓ : miktarını ölçen cihazı içermelidir. Bu durumda, doğru düzenek A şıkkı gibi olmalıdır. KOH çözeltisi, karbondioksiti

Message ChatGPT...

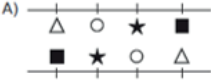
ChatGPT can make mistakes. Consider checking important information.


Figure 5. Question 3 for 2019 and ChatGPT's answer

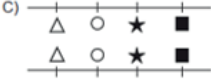
In this question, images of the experiment materials are provided, and the properties of the materials are detailed in the text. Additionally it is requested that the experiment setup be designed appropriately for its purpose using these materials. As seen in the answer in Figure 5, ChatGPT successfully answered the complex visual question. It successfully analyzed the experiment materials and correctly identified which materials were used to prepare the experimental setups in the options. Furthermore it correctly identified the correct option by explaining why materials appropriate for the purpose of the experiment should be used.

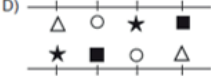
1. Bir öğretmen, öğrencilerinden Δ , \blacksquare , \circ , \star şekillerini kullanarak DNA modeli oluşturmalarını istiyor.

Buna göre öğrencilerin oluşturduğu aşağıdaki DNA modellerinden hangisi doğrudur?
(Zincirler üzerindeki şekiller nükleotitleri göstermektedir.)

A) 

B) 

C) 

D) 

ChatGPT 4

ChatGPT

Bu resimde, dört farklı DNA modeli gösteriliyor ve her bir model, nükleotitleri temsil eden dört farklı şekilden (üçgen, kare, yuvarlak, yıldız) oluşuyor. DNA'nın çift sarmal yapısını ve baz eşleşme kurallarını göz önünde bulundurarak, doğru modeli seçmemiz gerekiyor. DNA'da adenin (A) her zaman timin (T) ile eşleşir ve guanin (G) her zaman sitozin (C) ile eşleşir.

Model A, B, C ve D seçeneklerinde, eşleşmeler karşılıklı olarak görülüyor ve her çiftin simetrik olduğunu gözlemleyebiliriz. Bu nükleotit eşleşmesi kuralına göre, doğru modeli eşleşen şekillerle bir düzen oluşturması ve karşılıklı şekillerin her sarmalda tutarlı olması gerekmektedir.

Bununla birlikte, DNA'nın çift sarmal yapısını oluşturmak için sadece şekillerin eşleşmesine değil, aynı zamanda sarmallar arasındaki uzaklıkların da tutarlı olması gerekir, çünkü DNA'nın her dönüğünde aynı sayıda nükleotit bulunur. Bu detaylara göre, doğru modeli belirlememiz gerekmektedir. Bu bilgileri dikkate alarak seçenekleri değerlendirelim:

- Model A'da, her sarmal boyunca eşleşen şekiller yok, bu nedenle bu model doğru olamaz.
- Model B, tüm sıralarda karşılıklı gelen şekiller var ve her iki sarmal boyunca düzenli bir aralık gözlemleniyor. Bu, DNA'nın tutarlı genişlik ve eşleşme özelliklerine uygun.
- Model C'de, üstteki ve alttaki sıralar birbirini tutarlı bir şekilde takip etmiyor, yani bu model de doğru olamaz.
- Model D'de, üst sıra ve alt sıra arasındaki eşleşme düzenli değil, bu nedenle bu model de doğru olamaz.

Buna göre, sadece Model B, DNA'nın çift sarmal yapısını ve baz eşleşme kurallarını doğru bir şekilde temsil ediyor gibi görünüyor. Yani, öğrencilerden istenen DNA modeli, seçenekler içinde Model B olacaktır.

Figure 6. Question 1 for 2018 and ChatGPT's answer

As seen in the question and answer in Figure 6, ChatGPT was unable to successfully analyze the image. Upon examining the explanation, it is evident that the analysis performed by ChatGPT contains errors in the visually provided options.

When examining the answers given to visual-based questions, it was found that ChatGPT is generally successful in image processing. However, errors were observed, albeit at a low rate. Although the explanations provided to the questions are accurate and detailed, there are also instances where the answers are incorrect. This indicates that ChatGPT has a limited ability to effectively correlate the explanations it provides with the answers it ultimately delivers.

ChatGPT's Effectiveness in Text Based Questions

Figure 7 shows the distribution of ChatGPT's correct and incorrect answers to text based questions, as well as unanswered questions, between 2018 and 2023. In 2018, it answered all 2 text based questions correctly, while in 2019, it answered 4 out of 5 questions correctly and 1 incorrectly. In 2020, it answered all 5 questions correctly. In 2021, it answered 8 out of 9 questions correctly and 1 incorrectly, while in 2022, it answered all 6 questions correctly. In 2023, it was observed that it answered all 5 questions correctly. Out of a total of 32 text-based questions, it answered 30 correctly and 2 incorrectly,

leaving no questions unanswered. ChatGPT achieved a 94% accuracy rate on text-based questions, made a 6% error rate, and left no questions unanswered.

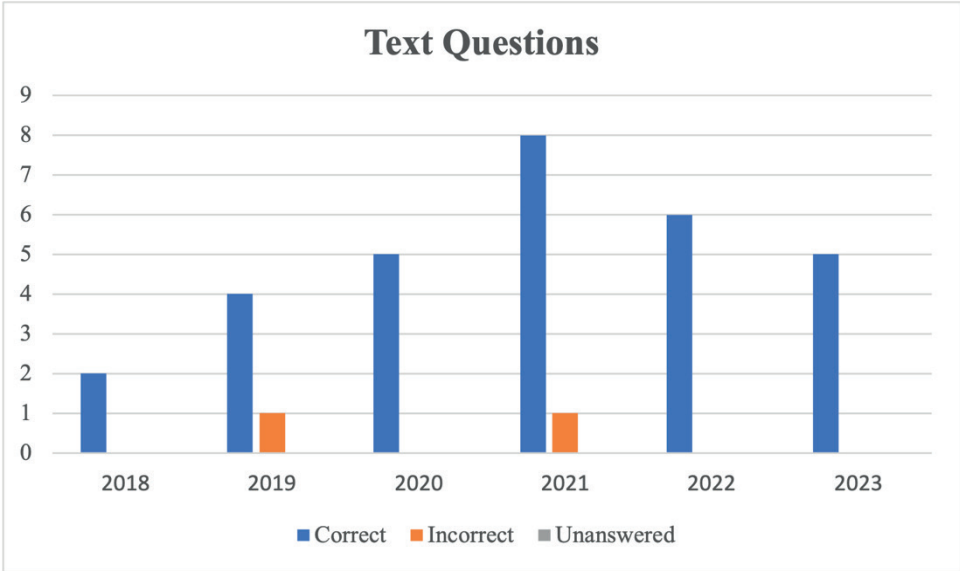


Figure 7. *Distribution of correct and incorrect answers to text based questions by year*

In Figure 8, ChatGPT was asked a text based question about how a heated substance changes depending on its type, mass, and temperature. The aim was to analyze this question and determine the correct option, evaluating its ability to interpret text based data and how it processes scientific concepts based on such data. In the answer to the question, it was observed that it correctly explained what was required for the desired experiment and successfully expressed the independent and constant variables in the experiment. However, when evaluating the options, it was found that it was unable to effectively relate the initial explanations to the options.

| | |
|---|---|
| <p>19. Saf bir maddenin sıcaklık değişimi ile madde miktarı arasındaki ilişkinin araştırıldığı bir deneyde; özdeş cam kaplara farklı miktarlarda, başlangıç sıcaklıkları eşit olan saf su ve saf etil alkol konuluyor. Su bulunan cam kap 5 dakika, etil alkol bulunan cam kap ise 10 dakika özdeş ısıtıcılarla ısıtılıyor. Isı alışverişi sadece sıvılar ve ısıtıcılar arasında gerçekleşiyor.</p> <p>Deneyin bu şekliyle araştırmanın amacına uygun olmadığı bilindiğine göre deneyde sıcaklık değişimi ve madde miktarı arasındaki ilişkinin doğru belirlenebilmesi için aşağıdaki işlemlerden hangisi yapılmalıdır?</p> | <p>ChatGPT</p> <p>Deneyin amacı, saf bir maddenin sıcaklık değişimi ile madde miktarı arasındaki ilişkiyi araştırmaksa, değişkenlerden yalnızca biri değiştirilmeli ve diğerleri sabit tutulmalıdır. Yani, madde miktarının sıcaklık değişimine olan etkisini incelemek için, her seferinde sadece madde miktarını değiştirip diğer koşulları (ısıtma süresi, başlangıç sıcaklığı, ısıtıcı gücü vb.) sabit tutulmalıdır.</p> <p>Bu deneyde su ve etil alkol gibi iki farklı madde kullanılıyor ve her bir madde farklı sürelerde ısıtılıyor. Bu durum, deneyi farklı değişkenlerin etkisini ölçmek için uygun hale getirmeye yardımcı olur ancak her iki madde de aynı ısıtma süresiyle ısıtılıyor. Bu durum, deneyi farklı değişkenlerin etkisini ölçmek için uygun hale getirmeye yardımcı olur ancak her iki madde de aynı ısıtma süresiyle ısıtılıyor.</p> <p>Seçenekleri değerlendirilim:</p> <p>A) Etil alkol bulunan kap 5 dakika süreyle ısıtılıp diğer değişkenler aynı bırakılmalı. (Bu durum, her iki sıvının da aynı süre ısıtılmasını sağlar, böylece ısıtma süresi değişkenini sabit tutarız, ancak hala iki farklı madde kullanıldığı için bu doğru bir düzeltme değildir.)</p> |
|---|---|

| | |
|---|---|
| <p>A) Etil alkol bulunan kap 5 dakika süreyle ısıtılıp diğer değişkenler aynı bırakılmalı.</p> <p>B) Su ve etil alkolün başlangıç sıcaklıkları farklı hâle getirilip diğer değişkenler aynı bırakılmalı.</p> <p>C) Cam kaplara sadece farklı miktarlarda etil alkol konulup kapların ısıtılma süreleri eşitlenmeli.</p> <p>D) Deney kaplarına sadece su konulup diğer değişkenler aynı bırakılmalı.</p> | <p>B) Su ve etil alkolün başlangıç sıcaklıkları farklı hâle getirilip diğer değişkenler aynı bırakılmalı. (Bu, deneyin amacına uygun değil çünkü başlangıç sıcaklıklarını değiştirmek başka bir değişkeni daha deney ekler.)</p> <p>C) Cam kaplara sadece farklı miktarlarda etil alkol konulup kapların ısıtılma süreleri eşitlenmeli. (Bu, aynı maddeyi kullanarak madde miktarının sıcaklık değişimi üzerindeki etkisini incelemek için daha uygun bir yöntem olur, ancak sıvı olarak sadece etil alkol kullanılır.)</p> <p>D) Deney kaplarına sadece su konulup diğer değişkenler aynı bırakılmalı. (Bu seçenek, deneyin saf bir maddenin miktarının sıcaklık değişimi üzerindeki etkisini incelemek için en uygundur çünkü sadece bir tür madde kullanılır ve diğer değişkenler sabit tutulur.)</p> <p>Doğru deney tasarımı için, madde miktarının etkisini ölçmek istiyorsak, aynı maddeyi kullanmalı ve diğer tüm koşulları (ısıtma süresi, başlangıç sıcaklığı, ısıtıcı gücü vb.) sabit tutulmalıdır. Bu nedenle doğru cevap "D" seçeneğidir.</p> |
|---|---|

Figure 8. Question 19 for 2019 and ChatGPT's answer

When examining the answers provided to all text based questions, ChatGPT was found to be generally successful. However errors were observed, albeit at a very low rate. Although the explanations it provided to questions were accurate and detailed, in some cases, its answers were found to be incorrect. This indicates that it may encounter limitations under certain scenarios and conditions and may not produce completely accurate results in every situation. While demonstrating that ChatGPT's information processing and integration capabilities may encounter limitations, it also shows that its ability to effectively relate the explanations it provides to the answers it ultimately delivers is limited.

ChatGPT's Effectiveness in Questions

The question asks about the applications of electrostatic charge in technology and how the repulsive and attractive forces between electric

charges function, presented in a text based format (Figure 9). When examining ChatGPT's response, while the conceptual information is generally correct, there are some instances where this knowledge falls short in practical application. In this question, it explains that if the paint is negatively charged, it will distribute evenly due to the attraction of opposite charges. However, in reality, negatively charged paint particles repel each other, resulting in a homogeneous distribution. This situation reveals that although ChatGPT has in-depth knowledge on certain topics, there may be limitations in its ability to use this knowledge correctly and effectively within a given context.

| | |
|---|---|
| <p>9. Genç bir girişimci olan Ali, beyaz eşyaların metal yüzeylerini boyama işine giriyor. Öğrendiklerini uygulamak isteyen Ali ilk denemede negatif (-) yükle yüklediği beyaz eşyanın metal yüzeyine nötr boya tanecikleri püskürttüğünde boyanın bu metalin yüzeyine düzgün dağılmadığını, bazı bölgelerde koyu ve açık renklerin oluştuğunu görüyor.</p> <p>Bu sorunu çözmek için Ali ikinci denemede, birinci denemedeyle özdeş olan metal bir yüzeyi yine negatif (-) yükle yükleyerek metalin yüzeyine bu kez pozitif (+) yüklü boya taneciklerini püskürttüğünde boyanın yüzeye düzgün dağıldığını görüyor.</p> <p>Buna göre ikinci denemede beyaz eşyanın metal yüzeyine boyanın düzgün dağılmasının sebebi aşağıdakilerden hangisi ile açıklanır?</p> <p>A) Nötr cisimlerin yüklü cisimler tarafından çekilmesi B) Zıt yüklü cisimlerin birbirini çekmesi C) Aynı yüklü cisimlerin birbirini itmesi D) Nötr bir cismin başka bir nötr cisim tarafından etkilenmemesi</p> | <p>ChatGPT</p> <p>İkinci denemede Ali'nin negatif yüklü metal yüzeye pozitif yüklü boya taneciklerini püskürttüğünde boyanın düzgün bir şekilde dağıldığını gözlemlemesi, elektrostatik ilkelere göre zıt yüklü cisimlerin birbirini çekmesi prensibiyle açıklanabilir. Ali'nin ilk denemede kullandığı nötr boya tanecikleri yüksüz olduğu için negatif yüklü metal yüzeye düzgün bir şekilde yapışmamıştır. Ancak ikinci denemede pozitif yüklü boya tanecikleri, negatif yüklü metal yüzeye elektrostatik çekimle yapışmış ve böylece düzgün bir dağılım sağlamıştır.</p> <p>Bu nedenle doğru cevap:</p> <p>B) Zıt yüklü cisimlerin birbirini çekmesidir.</p> <p><input type="checkbox"/> <input checked="" type="radio"/> <input type="checkbox"/> <input type="checkbox"/></p> |
|---|---|

Figure 9. Question 9 for 2019 and ChatGPT's answer

Appropriateness of ChatGPT Responses for Student Grade Levels

In this study, while examining ChatGPT's question solutions, the suitability of its answers for the student level was also taken into consideration.

In the question regarding the concept of pressure shown in Figure 10, the variables affecting solid pressure are presented both visually and in text form. ChatGPT was evaluated to determine whether it could provide a middle school level appropriate answer to this question. Upon examining the solution to the question, it was determined that the information provided by ChatGPT was not appropriate for middle school level. This is because the current middle school science curriculum states that mathematical connections are not covered in pressure topics. However, the answer uses mathematical expressions, making it unsuitable for the student's level. Despite this, ChatGPT provides accurate information in its analysis of the options.

11. Bir odada bulunan şekildeki ranzanın ayakları, üzerinde bulunduğu halı zeminde derin bir iz bırakmıştır.

Buna göre aşağıda verilenlerden hangisi yapılırsa ranza ayaklarının halı zeminde bıraktığı izin derinliği artar?

A) Ranzanın birinci katındaki ayaklarının altına genişliği, ayakların genişliğinden daha fazla olan bir tahta parçası koyulursa
 B) Ranzanın ikinci katında bulunan yatak daha ağır olanı ile değiştirilirse
 C) Ranzanın birinci katında bulunan yatak bulunduğu yerden çıkarılırsa
 D) Ranzanın ikinci katı alınıp başka bir yere koyulursa

ChatGPT
 Ranza ayaklarının halı zeminde bıraktığı izin derinliği, ranzanın uyguladığı basınçla ilgilidir. Basınç, bir kuvvetin birim alana uygulandığı zaman ortaya çıkar ve formülü $P = \frac{F}{A}$ şeklindedir, burada P basıncı, F kuvveti ve A alanı temsil eder. Eğer kuvvet aynı kalırken alan azalır, basınç artar ve halıda daha derin bir iz bırakır. Eğer kuvvet artar ve alan aynı kalırsa, o zaman da basınç artar.

Seçenekleri değerlendirelim:

A) Eğer ranzanın ayaklarının altına, ayakların genişliğinden daha fazla olan bir tahta parçası koyulursa, uygulanan kuvvet aynı kalırken alan artar. Bu durumda, basınç azalır ve halı üzerindeki izin derinliği azalır.

B) Ranzanın ikinci katında bulunan yatak daha ağır olanı ile değiştirilirse, uygulanan kuvvet artar. Alan değişmediğine göre, basınç artar ve halidaki iz daha derin olur.

C) Eğer ranzanın birinci katındaki yatak çıkarılırsa, uygulanan toplam kuvvet azalır. Bu durumda basınç azalır ve halı üzerindeki izin derinliği azalır.

D) Ranzanın ikinci katı alınıp başka bir yere koyulursa, ranzanın toplam ağırlığı azalır, bu da halı üzerindeki toplam basıncın azalmasına neden olur. Bu durumda izin derinliği azalır.

Bu bilgilere göre, ranza ayaklarının halı üzerinde bıraktığı izin derinliğini artıracak olan seçenek ikinci katta bulunan yatağın daha ağır olanı ile değiştirilmesi durumudur, yani doğru cevap B seçeneğidir.

Figure 10. Question 11 for 2020 and ChatGPT's answer

Comparison of HSTS Science Türkiye Average with ChatGPT

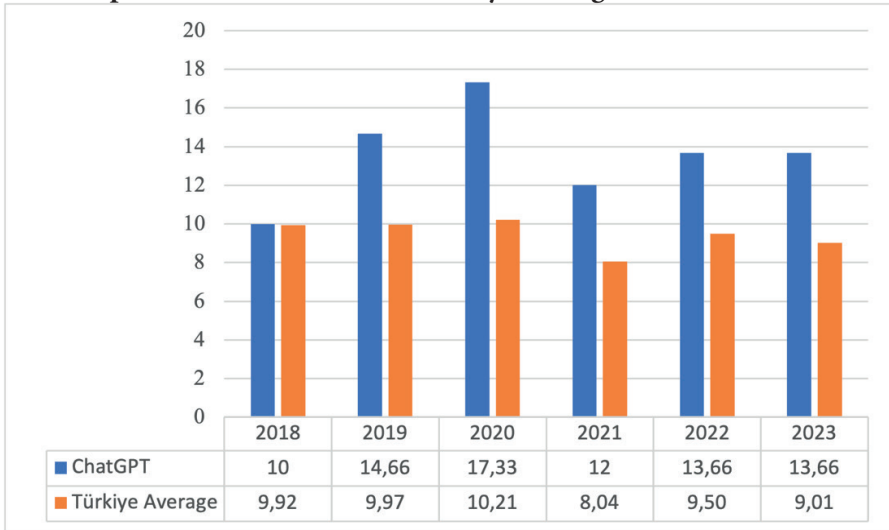


Figure 11. ChatGPT and Türkiye average net scores by year

The comparison of ChatGPT's scores and Türkiye's average net scores in HSTS science questions over the years is presented in Figure 11. In the study conducted on the solutions to HSTS science questions, ChatGPT's net scores over the years were compared with the averages across Türkiye. In 2018, ChatGPT's net score was 10, while the Türkiye average was 9.92, and these

values were found to be close to each other. In 2019, ChatGPT's net score rose to 14.66, while the Türkiye average was 9.97. In 2020, the chatbot achieved its highest net score of 17.33, while the Türkiye average was 10.21. In 2021, a significant drop was observed in the Türkiye average, and ChatGPT's net score remained at 12. In 2022, ChatGPT's net score rose to 13.66, while the Türkiye average was 9.5 net. In 2023, ChatGPT maintained its 2022 net score of 13.66, while the Türkiye average was 9.01 net, below the previous year's level. According to this data, ChatGPT's performance has consistently been above the Türkiye average and has shown overall stability. It has demonstrated particularly strong performance in 2020 and 2023.

Conclusion

The findings related to the first research question “What is the accuracy rate of the answers provided by ChatGPT-4 to HSTS science exam questions, and in which types of questions do incorrect responses most commonly occur?” indicate that the ChatGPT performed successfully, achieving an accuracy rate of 94% on text based questions and 69% on visually based questions. AI technologies also play a significant role in higher education by identifying students' strengths and weaknesses and providing automated feedback (Taşçı & Çelebi, 2020). However, these findings suggest that AI performs better on certain types of questions. While the high accuracy in text based questions reflects its advanced language processing capabilities, the lower performance in visually based questions points to a need for improvement in visual information processing and interpretation. Considering the importance of visual learning in science education, this highlights the necessity of further developing AI applications in educational contexts.

The findings for the second research question “To what extent is ChatGPT-4 capable of understanding HSTS science question texts and extracting the necessary information?” demonstrate that the ChatGPT is highly successful in understanding, organizing, and interpreting textual information. For instance, it can accurately distinguish between physical and chemical changes described in texts and effectively identify variables in experimental contexts. These results emphasize the strength of AI's language processing and analytical capabilities and its potential applications in education. With an error rate of only 6% in text based questions, AI shows a high level of accuracy and reliability (Kurtboğan, 2023). This success points to considerable potential for AI in educational technologies. When HSTS science results are examined through the lens of didactic transposition theory, the need for a better understanding of instructional strategies and student needs becomes evident (Taşkiran & Ergun, 2024). AI's advanced language processing capacity may provide a significant advantage by offering personalized learning materials and supporting individual learning processes, particularly in crowded classrooms, where it can reduce teachers' workload

and enable more individualized feedback (Taşkıran & Ergun, 2023).

The results addressing the third research question “How successful is ChatGPT-4 in processing and interpreting visual elements included in HSTS science questions?” show that the ChatGPT’s performance in visual processing and interpretation is not as strong as its text processing abilities. The error rate for visually based questions was found to be 26%, with an additional 5% of questions left unanswered. Overall, 31% of the responses to visually based questions were either incorrect or missing, indicating that AI encounters notable difficulties with visual content. For example, in complex experimental questions, the ChatGPT can more successfully interpret visuals involving detailed experimental setups; however, it may struggle with questions containing limited visual detail, such as partial representations of the periodic table. This suggests that the ChatGPT’s ability to differentiate and interpret improves as the richness of visual detail increases. Enhancing and diversifying visual datasets could improve AI performance in this area. Given the importance of visual information in education, improvements in this domain could significantly enhance AI’s overall educational effectiveness.

The findings for the fourth research question “How effective and detailed are the explanations provided by ChatGPT-4 for its answers to HSTS science questions?” reveal that the ChatGPT performs quite successfully in this regard. The results show that the ChatGPT typically begins by analyzing the question, explaining what is being asked, and presenting the necessary conceptual knowledge along with relevant definitions. For example, in a question about a DNA model, it explains which nucleotide base pairs with adenine and then evaluates the options to determine the correct answer. This demonstrates the AI’s ability to provide clear and detailed explanations. The findings highlight ChatGPT’s potential in education, particularly in enhancing students’ conceptual understanding. Rather than simply providing correct answers, AI explains the reasoning behind them, promoting deeper learning and supporting the development of students’ analytical thinking skills. This feature can help students better understand complex concepts and increase their confidence.

The findings related to the fifth research question “To what degree are ChatGPT-4’s responses to HSTS science questions appropriate for the students’ grade level?” indicate that the AI generally provides responses appropriate to students’ levels. However, in some cases, ChatGPT does not fully take students’ learning levels into account. This issue is particularly evident in physics related questions. For instance, although the 8th-grade science curriculum does not require establishing mathematical relationships for concepts such as the relationship between pressure and surface area, the AI was observed to provide mathematically based explanations in such cases.

These findings suggest that ChatGPT may sometimes challenge students' conceptual understanding and that greater attention should be given to aligning responses with students' levels. For AI to be effective in education, it is crucial that its responses match students' age and knowledge levels. This alignment supports students in benefiting fully from AI and enhances the efficiency of their learning processes.

Overall, the findings indicate that ChatGPT-4 performs successfully in answering HSTS science questions. Its high accuracy in text based questions demonstrates its effectiveness and consistency. This success suggests that AI can play an important role in education by assessing students' knowledge, identifying learning gaps, and providing automated feedback. However, the lower performance in visually based questions and occasional mismatches with student level highlight areas that require further improvement.

Recommendations

Based on the findings of the study, the following recommendations are proposed:

Further studies can be conducted using exams at different grade levels to examine the effects of ChatGPT-4 on science education in greater detail.

The performance of other large language models can be evaluated on science exam questions.

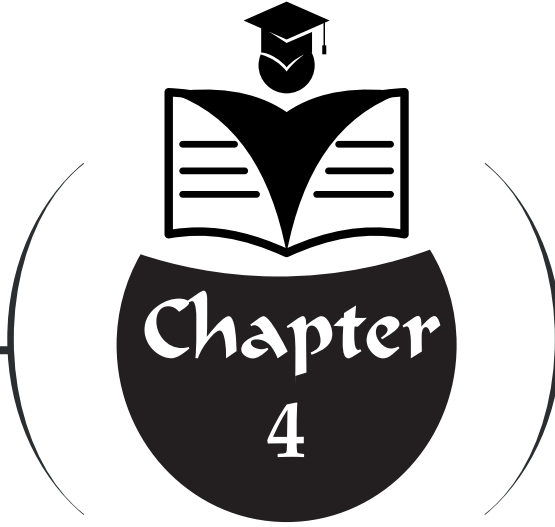
The problem solving approaches of science teachers in national exams can be compared with those of artificial intelligence models.

References

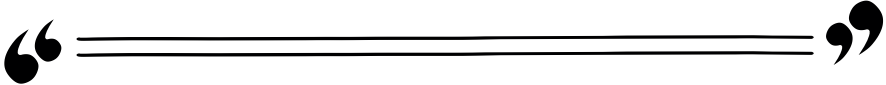
- Alessi, M. R., Gomes, H. C., Castro, M., & Okamoto, C. T. (2024). Performance of ChatGPT in solving questions from the progress test (Brazilian national medical exam): a potential artificial intelligence tool in medical practice. *Cureus*, <https://doi.org/10.7759/cureus.64924>
- Alfertschofer, M., Hoch, C. C., Funk, P. F., Hollmann, K. S., Wollenberg, B., Knoedler, S., & Knoedler, L. (2023). Sailing the seven seas: a multinational comparison of ChatGPT's performance on medical licensing examinations. *Annals of Biomedical Engineering*. <https://doi.org/10.1007/s10439-023-03338-3>
- Arvidsson, R., Gunnarsson, R., Entezarjou, A., Sundemo, D., & Wikberg, C. (2024). ChatGPT (GPT-4) versus doctors on complex cases of the Swedish family medicine specialist examination: an observational comparative study. *BMJ open*, *14*(12), e086148. <https://10.1136/bmjopen-2024-086148>
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M.T., Zhang, Y. (2023). Sparks of artificial general intelligence: early experiments with GPT-4. <https://doi.org/10.48550/arXiv.2303.12712>
- Dao, X. Q., & Le, N. B. (2023). Investigating the effectiveness of ChatGPT in mathematical reasoning and problem solving: Evidence from the Vietnamese national high school graduation examination. *arXiv preprint arXiv:2306.06331*. <https://doi.org/10.48550/arXiv.2306.06331>
- Flores-Cohaila, J., García-Vicente, A., Vizcarra-Jiménez, S., De La Cruz-Galán, J., Gutiérrez-Arratia, J., Torres, B., & Taype-Rondan, A. (2023). Performance of ChatGPT on the peruvian national licensing medical examination: cross-sectional study. *JMIR Medical Education*, *9*. <https://doi.org/10.2196/48039>
- Frieder, S., Pinchetti, L., Chevalier, C., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., ... & Berner, J. (2023). Mathematical capabilities of ChatGPT. *Advances in neural information processing systems*, *36*, 27699-27744. <https://doi.org/10.48550/arXiv.2301.13867>
- Gandhi, A., Joesph, M., Rajagopal, M., Aparnavi, M., Katkuri, M., Dayama, M., Satapathy, M., Mahalaqua, P., Khatib, N., Gaidhane, M., Quazi, M., Zahiruddin, S., & Behera, M. (2024). Performance of ChatGPT on the India undergraduate community medicine examination: cross-sectional study. *JMIR Formative Research*, *8*. <https://doi.org/10.2196/49964>
- Giannos, P., & Delardas, O. (2023). Performance of ChatGPT on UK Standardized admission tests: Insights from the BMAT, TMUA, LNAT, and TSA examinations. *JMIR Medical Education*, *9*. <https://doi.org/10.2196/47737>
- Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R., & Chartash, D. (2023). How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, *9*. <https://doi.org/10.2196/45312>

- Guigue, P. A., Meyer, R., Thivolle-Lioux, G., Brezinov, Y., & Levin, G. (2024). Performance of ChatGPT in french language parcours d'accès spécifique santé test and in OBGYN. *International Journal of Gynecology & Obstetrics*, 164(3), 959-963. <https://doi.org/10.1002/ijgo.15083>
- Hsieh, C. H., Hsieh, H. Y., & Lin, H. P. (2024). Evaluating the performance of ChatGPT-3.5 and ChatGPT-4 on the Taiwan plastic surgery board examination. *Heliyon*, 10(14), e34851. <https://doi.org/10.1016/j.heliyon.2024.e34851>
- Huang, C. H., Hsiao, H. J., Yeh, P. C., Wu, K. C., & Kao, C. H. (2024). Performance of ChatGPT on Stage 1 of the Taiwanese medical licensing exam. *Digital Health*, 10, 20552076241233144. <https://doi.org/10.1177/20552076241233144>
- Jin, H. K., Lee, H. E., & Kim, E. (2024). Performance of ChatGPT-3.5 and GPT-4 in national licensing examinations for medicine, pharmacy, dentistry, and nursing: a systematic review and meta-analysis. *BMC Medical Education*, 24(1), 1013. <https://doi.org/10.1186/s12909-024-05944-8>
- Kurtboğan, H. (2023). *Yeni dünyada yapay zeka metaforu ve yapay zekânın çalışan performansına etkisi* [Yüksek Lisans Tezi, Karamanoğlu Mehmetbey Üniversitesi, Sosyal Bilimler Enstitüsü, Yönetim Bilişim Sistemleri Ana Bilim Dalı]. Karaman
- Lu, J., Zheng, R., Gong, Z. & Xu, H. (2024) Supporting teachers' professional development with generative ai: the effects on higher order thinking and self-efficacy, *EEE Transactions on Learning Technologies*, 17, 1267-1277. <https://doi.org/10.1109/TLT.2024.3369690>
- Ming, S., Guo, Q., Cheng, W., & Lei, B. (2024). Influence of model evolution and system roles on ChatGPT's performance in Chinese medical licensing exams: comparative study. *JMIR Medical Education*, 10(1), e52784.
- MoNE, (2025) The High School Transition System. <https://odsgm.meb.gov.tr/www/liselere-gecis-sistemi/icerik/1012>
- Oztermeli, A.D. & Oztermeli, A. (2023). ChatGPT performance in the medical specialty exam: An observational study. *Medicine* 102(32):p e34673, <https://doi.org/10.1097/MD.00000000000034673>
- Rossettini, G., Rodeghiero, L., Corradi, F., Cook, C., Pillastrini, P., Turolla, A., ... & Palese, A. (2024). Comparative accuracy of ChatGPT-4, Microsoft Copilot and Google Gemini in the Italian entrance test for healthcare sciences degrees: a cross-sectional study. *BMC Medical Education*, 24(1), 694.
- Sawamura, S., Kohiyama, K., Takenaka, T., Sera, T., Inoue, T., & Nagai, T. (2024). performance of chatgpt 4.0 on japan's national physical therapist examination: a comprehensive analysis of text and visual question handling. *Cureus*, 16. <https://doi.org/10.7759/cureus.67347>
- Siebielec, J., Ordak, M., Oskroba, A., Dworakowska, A., & Bujalska-Zadrozny, M. (2024). Assessment Study of ChatGPT-3.5's Performance on the final Polish medical examination: accuracy in answering 980 questions. *Healthcare*, 12(16), 1637. <https://doi.org/10.3390/healthcare12161637>

- Solmonovich, R. L., Kouba, I., Quezada, O., Rodriguez-Ayala, G., Rojas, V., Bonilla, K., ... & Bracero, L. A. (2024). Artificial intelligence generates proficient Spanish obstetrics and gynecology counseling templates. *AJOG Global Reports*, 4(4), 100400.
- Tan, E. J. L., Ramos, K. A. L., Nazario, M. E. K. B., Lim, S. V. D., & Chu, S. B. (2024). AI to the test: measuring ChatGPT's objective accuracy in the SATs in comparison to human performance. In *2024 IEEE 48th Annual Computers, Software, and Applications Conference COMPSAC*, 153-158.
- Taşçı, G., ve Çelebi, M. (2020). Eğitimde yeni bir paradigma: Yükseköğretimde yapay zeka. *Uluslararası Toplum Araştırmaları Dergisi-International Journal of Society Researches*, 16(29).
- Taşkıran, B. & Ergun, M. (2023). *Fen eğitiminde yapay zeka destekli sistemlerin öğrenci takibinde kullanımının incelenmesi: Derspektif. III*. International Artificial Intelligence and Data Science Congress, İzmir Kâtip Çelebi University, İzmir, Türkiye.
- Taşkıran, B. & Ergun, M. (2024). *LGS Fen bilimleri raporuna didaktiksel dönüşüm teorisi çerçevesinden bakış. 3*. International Cankaya Scientific Studies Congress, (ss.608-616). Ankara, Türkiye.
- Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: insights from a meta-analysis. *Humanities and Social Sciences Communications*, 12(1), 1-21. <https://doi.org/10.1057/s41599-025-04787-y>
- Yıldırım, A., ve Şimşek, H. (2021). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin Yayıncılık.
- Zhong, W., Cui, R., Guo, Y., Liang, Y., Lu, S., Wang, Y., Saied, A., Chen, W. & Duan, N. (202). Agieval: A human-centric benchmark for evaluating foundation models. In *Findings of the association for computational linguistics: NAACL 2024*, 2299-2314.
- Zong, H., Li, J., Wu, E., Wu, R., Lu, J., & Shen, B. (2023). Performance of ChatGPT on Chinese national medical licensing examinations: a five-year examination evaluation study for physicians, pharmacists and nurses. *BMC Medical Education*, 24. <https://doi.org/10.1101/2023.07.09.23292415>



**MAPPING WHAT
RESEARCH CANNOT SAY:
METHODOLOGICAL REFLECTIONS
ON CROSS-NATIONAL
BIBLIOMETRIC ANALYSIS OF
DOCTORAL DISSERTATIONS**



Bora BAŞARAN¹

¹ Assoc. Prof. Dr., Anadolu University, Faculty of Education, Department of Foreign Language Education, bbasran@anadolu.edu.tr, Orcid: 0000-0003-0251-5895

All research methodologies have a dark side; a section of the empiric world that is systematically missed by the methodology's design. The survey instrument misses what respondents cannot, or will not, express verbally. Ethnographic studies miss what is not visible at the field site. Experimental studies miss the long term, diffused and structurally embedded effects that cannot be isolated even under the most carefully controlled conditions. The bibliography of all methodological approaches is, in fact, largely a bibliography of acknowledged shadows – acknowledging what the instrument reveals and what it obscures.

Comparative Education Research, as a field, is still developing its methodological traditions. One of those emerging traditions is the use of bibliometric analyses of academic dissertations. The potential of this approach is substantial: the systematic, replicable and large scale mapping of what doctoral students study, within which discipline, using what theoretical language and over time. However, the shadow cast by this approach has been subject to limited analysis. Those studies using bibliometric approaches generally focus their methodological discussion on issues related to the selection of sources of data, the removal of errors and the use of analytical software. These discussions treat the method itself as being technically difficult but epistemologically uncomplicated. The purpose of this Chapter is to argue that this treatment of the method is insufficient and that the epistemological implications of conducting cross national bibliometric analyses of doctoral research require separate consideration.

The arguments presented are based on a specific case: a comparative bibliometric analysis of 645 doctoral dissertations completed in Turkey and Germany between 2019 and 2024, where each dissertation was analyzed across five dimensions: the location of the dissertation in terms of disciplinary focus, the clustering of thematic content, the theoretical and methodological orientations of the dissertation, the temporal trends in the dissertation, and the co occurrence of policy vocabulary with other content areas (Başaran, 2025). Although this study was conducted with high levels of methodological transparency, including the acknowledgment of its dependence on title-keywords-abstracts indicators rather than full texts, this transparency provides a useful point of reflection on what, precisely, the method mapped and what it was structurally unable to map. This chapter is not intended to criticize the study in question but to provide a reflective surface; a worked example, from which six fundamental structural tensions inherent in the use of cross-national bibliometric research can be identified, named, and addressed.

Bibliometrics in Comparative Education: The Documentary Turn and Its Discontents

In dealing with the “levels of analysis” issue ; the methodological difficulty of comparing phenomena which are located in, and are formed by, qualitatively different national, cultural, and institutional contexts ; comparative education has grappled for years with the problem described by Bray and Thomas (1995). Although the classic debate of the field ; concerning whether there exist generalizable laws that can be developed from systematic comparative data versus the view that each nation’s educational system is uniquely particular and therefore cannot be compared ; has not been settled in terms of theory, the question of how to deal with the problem has been institutionally solved through an increase in the number of mixed-method and multi-level methods (Bray et al., 2014).

Documentary turns in educational research have also provided access to large scale comparisons of the same quality as traditional surveys and ethnographies. Documentary turns involve the systematic analysis of texts (policy documents, curriculum documents, textbooks, etc.) as data instead of as only context(Sarkiler, 2025a, 2025b). This has enabled a type of large-scale comparison that does not require the expensive, time-consuming data collection process associated with both surveys and ethnographies. The increased accessibility of bibliometric analysis to educational researchers lacking formal training in informetrics has been due largely to the development of Aria and Cuccurullo (2017) R-package bibliometrix and the spread of VOSviewer as a tool for co-citation and keyword network mapping.

In addition to the many studies examining what scholars study and cite, educational bibliometric studies have produced a relatively sparse literature regarding what these studies cannot show. While educational bibliometric studies follow a recognizable format of describing the source of their data, reporting keyword frequencies, discipline distributions, and co-citation networks, identifying growing areas of interest and waning areas of concern, and using bibliometric results to suggest gaps in research or areas where policymakers should focus their efforts, the assumptions underlying the methodology (i.e. what form of knowledge do the maps represent; what types of scholarly activities do the maps miss; how does national context influence the relationship between the map and the territory?) are rarely addressed directly within the studies themselves.

Indicator-Level Epistemology: What Bibliometrics Actually Measures

The term “indicator-level epistemology” is used to describe a particular epistemological condition of the researcher, i.e. the condition of accessing a complex, multi-dimensional phenomenon (the topic of a doctoral dissertation, the scope of a research field, a national academic culture) through a structured

set of indicators (titles, keywords, abstracts, classification codes) that were originally developed to be an administrative and retrieval tool for library systems rather than a research instrument. The condition described above differs structurally from the limited access to subjective meaning provided by a survey instrument (in this case, the questionnaire designer determines what will be asked), or the selective access to observable behavior available in ethnographic field work (in this case, the researcher determines which events to attend).

In bibliometric analysis, the instrument used to produce the data is not the product of the researcher but rather the metadata structure of national library systems, which were created for discovery, cataloging and archival purposes rather than for systematic thematic analysis. This results in a very important consequence: the data in a bibliometric study is based on the information contained in the metadata record for each dissertation, not the information actually contained in the dissertation itself. The metadata record reflects several things including:

1. The conventions of the national library system that generated the metadata record. (Example: the Deutsche Nationalbibliothek uses a Dewey Decimal system to categorize subjects while YÖK has its own unique taxonomy.)
2. Keyword practices of individual doctoral students and their advisors.
3. Abstracting practices of different disciplines.
4. Institutional incentives to increase the visibility and retrievability of research.

None of the above represents a neutral window into the intellectual content of dissertations. Maps of metadata represent maps of metadata, not maps of knowledge. The relationship between the two is not always consistent or stable and is, therefore, usually invisible to readers and researchers unless there is an explicit consideration of methodology.

While acknowledging the existence of indicator-level epistemology as an appropriate theoretical context for understanding cross-national bibliometric studies is not a rejection of the utility of such studies. Maps of metadata can be useful because they show the patterns of labeling of scholarship, the patterns of classification of institutions and the patterns of self-representation by discipline. These patterns are social facts about academic fields and are meaningful in and of themselves. However, the same cannot be said for over-interpreting the results of such studies (i.e. mistaking the map for the territory). Each of the six structural tensions discussed below represent an example of the potential for over-interpretation.

Structural Tension I: The Metadata Dependency Problem

The Gap Between Signal and Content

The most basic structural tension in indicator-level bibliometric research is the gap between what a dissertation's metadata indicates, and what the dissertation really has. The title encodes the author's ;and usually the supervisor's, interpretation of the most relevant to searchers, appropriate for their discipline, and effectively rhetorical description of the work. Keywords encode the author's interpretation of the vocabulary that they believe will best link the work with its desired audience. An abstract encodes a highly condensed representation of a study's goals, methods and results which can be subject to institutional review restrictions when they were written under strict word counts.

Each encoding includes systematic distortion. Titles usually do not indicate theoretical frameworks: a dissertation titled "Family Dynamics and Academic Achievement in Secondary Schools" could have been developed using a variety of theoretical frameworks such as Bourdieu's theory of cultural capital, ecological systems theory, or a cognitive psychology model of parental involvement ,the title does not make this distinction. Keywords usually are less current in terms of the use of keywords within the discipline: researchers who are trained in disciplines with well-established keyword usage (e.g., clinical psychology, economics) will develop more standardized keyword sets compared to researchers from inter- or transdisciplinary areas or new areas of study; therefore creating artificial cross-disciplinary differences. Abstracts, especially in areas of research where abstracts are completed before the dissertation has been finalized (as occurs in some German universities), may not represent the author's thesis argument at the time it was completed.

The German-Turkish Case

The data dependence issue revealed itself in multiple analytically interesting manners in the comparative dissertation study upon which this chapter is based. Dissertations were theoretically framed using title-keyword-abstract scanning to determine their theoretical framework classifications ; for example ecological/systems theory, feminist theory, sociocultural approach etc.

As a result, there were large numbers of dissertations classified into an unknown ('bilinmiyor') category within the Turkish corpus and relatively few dissertations had explicit theoretical frameworks labelled in the German corpus; both instances reflect the conventions in which these dissertations were documented and submitted, rather than the actual nature of the theoretical landscapes within which the respective bodies of research were undertaken. In terms of the cultural influences on German academia, its

historical development has produced a style in which theoretical frameworks are almost always developed in the body of the theory chapter, and nearly never explicitly referenced in the abstract or title. As a result, the theoretical orientation of dissertations from German universities can only be inferred in relation to what appears in the *Theorieteil* (Theory Chapter), and therefore appear to be largely absent from analysis at the indicator level. This cultural embeddedness of meaning is not confined to academic writing; it manifests equally in literary production, as demonstrated in translation studies of Turkish-German literary texts where cultural concepts resist direct linguistic transfer across the two traditions (Arabacıoğlu, 2020).

Turkish doctoral culture has also been influenced by an increasing number of standardized submission requirements established by YÖK regarding the format of thesis submissions, including structured abstracts with separate methodology and theoretical framework sections. Therefore, while dissertations from Turkish universities have greater visibility at the metadata level, they may do so in a way that overestimates the diversity of methodologies employed (e.g. when ‘mixed methods’ as a keyword describes studies employing completely different actual combinations of methodologies), and underestimates the theoretical sophistication of the research being reported (i.e., since keyword based classifications do not enable an assessment of how the theoretical framework was actually employed in the study).

Response Strategy: The Full-Text Validation Sub-Sample

Full-Text Validation Sub-Sample: A methodology used as a means to deal with the issue of the Metadata Dependency Problem. A randomly selected sub-sample of dissertations for which full-text access is available and an additional level of investigation occurs. The results from this sub-sample are used to calibrate the interpretation of the indicators at the level of the indicator. In the German-Turkish comparative study, a sub-sample of 20 dissertations (10 from each country) were analyzed using full-text content. The sub-sample of 20 dissertation validated calibration data: providing evidence that the thematic clusters established through metadata corresponded, at least in terms of direction, to actual dissertation content; however, it also identified areas (e.g., theoretical framework and methods) where metadata did not fully represent the content.

However, 20 dissertations from 645 (approximately 3.1%) provide a very small sample for validating. Cross-national bibliometric studies should strive for validation sub-samples of at least 10-15% of all dissertations within the sample to insure that calibration captures variability among the dimensions most important to the study’s conclusions. The results of the validation analysis should also be reported separately and in sufficient detail so that readers can judge the size of the difference between the signal indicated at the indicator-level and the content of the full-text.

Structural Tension II: The Taxonomy Translation Problem

Incommensurable Classification Systems

Cross-national bibliometrics needs to translate different national classifications to use for cross-national comparisons. At first glance, this translation issue seems simple; however, national libraries and research classification systems represent different academic disciplines, institutional history, and theories of how to organize knowledge. Therefore, there can be no complete translation from one system to another, as each classification represents a unique “cut” or perspective through the intellectual landscape. The Deutsche Nationalbibliothek uses a Dewey Decimal Classification (DDC)-based subject coding system. The DDC was developed based upon an Anglo-American intellectual tradition from the 19th century. Therefore, the DDC assumes certain boundaries between academic disciplines and their respective weights are embedded in the system in ways that are not directly comparable with the academic tradition of Germany. The well-known 360 Cluster (Social Problems, Social Services, Insurance) in the DDC is an example of this. The Cluster combines three major academic fields: social work, social policy, family welfare and insurance studies, with criminology. There is no corresponding Fächergruppe in the German University System that corresponds to these five fields. In contrast, the YÖK system classifies dissertations by the Anabilim Dalı (Major Academic Department) in which they were produced. Thus, the YÖK system creates a classification that is based on the organizational structure of the institutions involved, and is thus dependent on the departmental structure of the institution producing the dissertation, rather than the actual subject matter of the dissertation.

Standardisation and Its Costs

The German-Turkish Study solved the Taxonomy Translation Problem via a Standardization Procedure: by mapping each system to a common discipline (sociology, education, psychology, social services, law, economics, etc.) using a Content-Matching Process administered by the Principal Investigator. The Principal Investigator was familiar with both German and Turkish Academic Cultures and Languages. This is a Methodologically Reasonable Approach and has been the Method of Choice in Most Cross-National Bibliometric Studies, however it involves Interpretative Choices that warrant Examination. Consider how the German DDC 360 Cluster was treated in the Study’s Reported Results. The 360 Cluster represented 13.3% of the German Corpus - the Second Largest Disciplinary Grouping. It appeared as a Separate Category in the Cross National Comparison Table, with a Footnote explaining it was based on the DDC. The Transparent Treatment of this Discipline was Commendable. However, the Interpreted Question Remains: Is the 360 Cluster a Discipline like Sociology (47.1%) or Psychology (0.7%), or is it merely

an Administrative Artifact of the DNB Catalog System - a “Catch-All” which Collectively Captures Heterogeneous Content that Individually Classified Would be Distributed Across Family Therapy, Public Health, Social Work and Social Policy?

If the 360 Cluster Contains Dissertations That Were Classed in Turkey Under Social Services, Psychology, or Education, Then the Apparently Stark Difference Between the Two Countries (German Sociology-Dominated; Turkish Multi-Disciplinary Dispersed) May Be Partially an Artefact of Classification System Architecture Rather Than a Genuine Difference in the Intellectual Organization of Family Research.

Response Strategy: Double Coding and Sensitivity Analysis

The best solution to the taxonomic translation problem is using both double coding and sensitivity analysis. Double coding means that each dissertation’s placement into the categories of the cross-nationally shared system will be reviewed for accuracy by two coders – preferably one from each of the countries involved in the project, to bring different sets of background assumptions about how to draw the boundaries of disciplines. Sensitivity analysis means that the researcher tests his/her main conclusions about the results of the study based on alternative plausible ways to categorize the data. For example, if the researcher concludes that there is a high degree of sociological concentration in Germany based on the use of the DDC 360 cluster as part of the sociology discipline, then he/she has demonstrated that the conclusion is robust. If the researchers conclusions depend heavily on the way they have chosen to categorize the DDC 360 cluster, then the conclusions require a more cautious presentation.

It would greatly enhance the evidence-based value of comparative bibliometric research if all cross-national bibliometric research reported on sensitivity analyses for the taxonomy translations used. It would especially be important for research that involves ambiguous taxonomies like DDC 360. Reporting on which conclusions are stable under alternative taxonomic translation schemes, and which are sensitive to the researcher’s choices regarding taxonomic translation would significantly add to the value of comparative bibliometric research.

Structural Tension III: The Visibility Asymmetry Problem

Explicit vs. Embedded Knowledge

The problem of bibliometric analysis’s visibility asymmetry refers to how indicator-level bibliometric analysis favors the explicit knowledge labelled in metadata over tacit knowledge and embedded knowledge or terminology variability in the doctoral dissertation. An example of this is when a dissertation names its theoretical framework as “Bronfenbrenner’s ecological

systems theory” in the abstract or keywords of the dissertation then it will be captured in the analysis of theoretical framing. If the same dissertation uses the same theoretical framework without naming it in the dissertation, it will not appear in that analysis. Another example is that of a dissertation with a central theme stated in the title as “migrant family integration” will be captured in the analysis of migration themes; whereas, a dissertation that examines migration dynamics in a chapter titled “Socioeconomic Contexts of Family Education” without using the word “migration” in the dissertation’s metadata will not be captured in that analysis.

This cross-national dimension of this visibility asymptote does not randomly occur across different cultures. While the practices of making knowledge explicit in bibliographic metadata vary substantially across different cultural environments, the way that this expression variety interacts with the way that the dissertations are presented on databases (i.e., through keywords) does so in a non-random way. These two ways of expressing expressed knowledge, i.e., culture and database presentation, have developed in a way that creates a systematic disparity in the representation of German doctoral dissertations in any indicator-level analysis that relies on these two formats for analysis. The systematic disparity is created by the fact that German doctoral dissertations generally employ a Humboldtian tradition of theoretical depth which results in theoretical commitments being embedded in the body of the dissertation rather than being flagged in metadata, resulting in a systematic underrepresentation of German doctoral dissertations’ theoretical sophistication in such analysis. This pattern stands in contrast to Turkish academic work in the field of German Studies (Germanistik), where theoretical and conceptual frameworks are typically made explicit in abstracts and keywords — as illustrated, for example, in studies on intercultural sensitivity and cultural stereotypes in Turkish-German contexts (Arabacıoğlu & Balkaya, 2023).

Disciplinary Visibility Gradients

The visibility asymmetry issue also has a second dimension, which is how different disciplines are compared both within and across countries. Disciplines that have a very strong quantitative tradition, like economics and experimental psychology, can easily transfer their vocabulary (statistical methodologies, measurement instruments, dependent and independent variables), into abstract and keyword fields. This results in more standardized, keyword-rich metadata being produced by these types of disciplines. Disciplines with a strong interpretive tradition, like sociology of culture and educational anthropology, will generate much more varied, and as such, less comparable metadata. In the case of the German-Turkish comparison, this disciplinary visibility gradient likely makes the “methodological diversity” of Turkish dissertation studies appear greater than it actually is when compared

to those of Germany. The larger proportion of quantitatively and qualitatively described or labeled designs (i.e., quantitative, qualitative, and mixed-method designations) in the Turkish data set does not necessarily indicate that the Turkish dissertation studies demonstrate a greater variety of methods in conducting their research. It merely indicates that they label their methods more clearly ; due to the structural nature of the thesis submission process required by YÖK, and not necessarily due to the actual practices employed by researchers using different methods.

Response Strategy: Language-Sensitive Thesaurus Expansion

The most effective solution to the visibility asymmetry issue is through the application of a language-sensitive thesaurus, which is a comprehensive, systematically prepared vocabulary list for a particular theoretical approach or area of research. This list will not only provide the standard or canonical label for a particular approach or area of research, but also include all of its synonyms; near-synonyms; disciplinary variants; and language-specific terminological alternatives. In developing a thesaurus for the German-Turkish comparison, this means developing, for each major analytical category, not only the English-language canonical term, but the German and Turkish equivalents, near-equivalents, and typical disciplinary formulations.

For instance, a thesaurus entry for the “ecological systems theory” category would contain not only “Bronfenbrenner” and “ecological systems,” but also the German terms “ökologische Systemtheorie,” “sozial-ökologischer Ansatz,” “Mehrebenenperspektive,” and common partial indicators such as “Mikro-Meso-Makroebene,” as well as Turkish equivalents “ekolojik sistemler kuramı,” “çevresel sistemler modeli,” and associated formulations. Utilizing such an expanded thesaurus in a second pass analysis, and reporting on the proportional change in classification rates produced by this analysis, provides a direct quantitative measure of the size of the visibility asymmetry in any single study.

Structural Tension IV: The Unit-of-Analysis Problem

The Dissertation as Dated Artifact

A bibliometric analysis of all doctoral dissertations is treated as individual units of scholarly output ,as an entry into data for constructing frequency tables, trend lines and tabulations. The need to treat each dissertation in such a way is operational but is also a theoretical problem because a dissertation is not merely a unit of scholarship. It is a document written during a particular time frame in a scholar’s career; it is influenced by the supervision of the scholar; it is required to meet certain institutional guidelines for both content and structure and is often revised over a multiyear writing process where the start and stop points of the writing process are not necessarily the same as the date of graduation or completion.

The date-of-completion code used in almost every dissertation database captures not when a research project began but when it was administratively finalized. This temporal difference between when research occurs and when a dissertation is complete causes systematic error in the interpretation of trends in dissertation research: a dissertation submitted in 2019 but completed in 2024 would be recorded as having been completed in 2024. However, the research design, choice of theoretical framework and empirical focus for the dissertation were likely developed in 2019–2020. For a study spanning 2019–2024; a period which included the significant methodological disruption of the COVID-19 pandemic and the rapid adoption of large language models in academic research (Arabacıoğlu & Tunçyüz, 2025), this temporal difference is not insignificant. This transformation is particularly visible in Turkish foreign language education, where the integration of Web 2.0 tools, mobile applications, and AI-based learning platforms has substantially altered both pedagogical practice and curricular priorities within the timeframe covered by this study (Yılmaz Güngör, 2025).

The Dissertation as Social Node

The second aspect of the Unit-Of-Analysis problem deals with how the dissertation relates to other academic products. As a dissertation is typically not an independent work, there are generally accompanying academic works including conference presentation papers, working papers, and/or published articles within a scholarly peer-reviewed journal. A new form of the dissertation that is becoming prevalent in the Social Sciences in Germany is the “cumulative dissertation” (kumulative Dissertation), or a thesis presented as a compilation of previously or concurrently published peer-reviewed journal articles. This is essentially eroding the distinction between the dissertation as a document versus the dissertation as a collection of published articles each with their own bibliometric life. This is relevant for comparative international studies as in general the cumulative dissertation format is more commonly used in Germany, while the monograph style dissertation is predominant in Turkey. For example, a German dissertation that is submitted as a cumulative thesis, can be comprised of three to four different journal articles detailing the dissertation’s theory, empirical study results, and methodology innovation(s). Therefore, the dissertation’s title, abstract, and keywords that comprise the dissertation’s metadata capture only the framework document, (klujtjige Rahmenschrift), and not the content that is present in the articles.

Response Strategy: Dissertation-Publication Linkage

The way you respond to the unit-of-analysis issue with methodology is to make an explicit structural decision on scope: when conducting cross-national bibliometric studies of doctoral research, do you choose to trace how dissertations continue into their publication afterlives; or do you choose

to recognize and accept the dissertation document as your unit of analysis, and acknowledge what this implies for your study? While each option has some merit, the decision must be made explicitly and its implication(s) acknowledged. A dissertation-publication linkage procedure ; linking the journal articles that constitute a cumulative thesis ; could significantly reduce the information loss mentioned above if cumulative thesis formats were prevalent in one of the countries being studied (e.g., Germany). A dissertation-publication linkage procedure can be implemented using DOI-based linking, and matching authors' names and dates. Implementing such a procedure will likely require additional data infrastructure; however, it would increase the comparable quality of German and Turkish dissertation data in any future replication of this study.

Structural Tension V: The Policy-Proximity Problem

Co-occurrence Is Not Causation

One of the most analytically powerful; and methodologically hazardous, procedures in the comparative dissertation study is the policy co-occurrence analysis: examining how frequently dissertation metadata contains terms that appear also in national educational policy documents. The finding that Turkish dissertations on value transmission show significantly higher co-occurrence with policy vocabulary than their German counterparts is interpreted as evidence that Turkish doctoral research on the family is more closely aligned with, and possibly more directly shaped by, national educational policy imperatives.

This interpretation is plausible; but co-occurrence analysis cannot establish it. The co-occurrence of a research theme and a policy term in the same document's metadata is consistent with at least four different causal stories: (a) the research was directly prompted by, or designed to respond to, the policy; (b) the research and the policy both respond independently to the same underlying social phenomenon; (c) the researcher adopted the policy's vocabulary for strategic purposes (discoverability, funding eligibility, institutional legitimacy) without the research itself being substantively policy-driven; and (d) the co-occurrence reflects the shared vocabulary of a broader public discourse in which both policy documents and doctoral research participate, without any direct relationship between them.

The Vocabulary Contamination Risk

A second dimension of the policy-proximity problem is what might be called vocabulary contamination: the risk that the policy vocabulary used to construct the co-occurrence search terms is itself so pervasive in a particular national discourse that it appears in dissertation metadata for reasons entirely unrelated to the research's substantive policy orientation. In Turkey,

terms like ‘değerler eğitimi’ (values education), ‘aile-okul işbirliği’ (family-school partnership), and ‘maarif’ appear extensively in doctoral dissertation metadata not necessarily because the researchers were responding to specific policies but because these terms have become part of the taken-for-granted vocabulary of Turkish educational discourse ; what Bernstein (1975) would call the ‘classification principles’ of the educational field itself.

When the policy co-occurrence analysis uses these terms as evidence of policy proximity, it risks systematically inflating the apparent policy-responsiveness of Turkish research relative to German research, where the corresponding policy terms (Wertevermittlung, Interkulturalität, Inklusion) are less pervasively present in dissertation metadata ; not because German doctoral research is less policy-relevant but because German academic culture maintains a sharper stylistic distinction between policy documents and scholarly discourse. This structural distinction is consistent with analyses of Turkish foreign language education more broadly, where curriculum-centric and examination-oriented policies have been shown to shape not only pedagogical practice but also the thematic priorities of educational research (Sarkiler, 2025c).

Response Strategy: Temporal Sequencing and Discursive Context Analysis

The main way to respond to the issue of proximity to policies, is by using time; instead of viewing policy terms appearing in both dissertation metadata and policy documents at the same time as proof of the policy terms being used in both places at the same time, researchers could analyze how quickly policy terms start to show up in dissertation metadata once they have appeared in policy documents.

If the dissertation topics related to the policy terms consistently follow significant policy announcements (as opposed to preceding them or showing up at the same time), then the “policies drive research” view becomes stronger.

If the dissertation topics are shown to have appeared at the same time, or even before the policy announcement, then it would make sense to believe in some type of common cause (i.e., a general idea or trend) or that there was a shared level of discourse among the research community which led to the use of policy terms in dissertations.

Temporal sequencing and discursive context analysis can take what would otherwise be a correlation between two sets of data (co-occurrence) and turn that into a relationship of a specific type (i.e. policy terms were used in dissertation topics either because of, or independently of the policy). Together, these two methods can turn the association between the two data sets into a set of very definable claims about how policy terms were interpreted and/or used in dissertations.

Structural Tension VI: The Generalisability Problem

From Corpus to National Academic Culture

The biggest leap of faith taken by researchers using cross-national bibliometrics is the jump from a set of dissertations to an inference about an entire nation's culture and/or paradigm(s) for research, or an entire nation's policy environment. A similar inference is made by comparing the DNB and YÖK corpora and identifying specific patterns within them as indicative of differences in how scholars create their knowledge of the family in each country's scholarly domain. These are large, important and actionable statements ; they form the basis for the policy recommendations included in companion work (Basaran, 2025), and the structural reform proposals that flow from them.

However, there are a series of inferential steps from the observed patterns in the corpora to the inferred national culture and generalizations. Each of these steps should be treated separately and explicitly. Firstly, the corpus is a sample ; although the study attempts to comprehensively include all possible dissertations through use of key word searching, the selection criteria (that the dissertation includes the terms “family,” “Familie” or “Aile” in either title, key words, or abstract) exclude dissertations where the family is a major theme but not the central theme. Secondly, the corpus is a snapshot ; this can range anywhere from 6 years of doctoral output; it does not necessarily reflect long term trends. Thirdly, the corpus is limited to a single type of scholarly output ; doctoral dissertations ; and the relationships among this type of scholarly output and the overall national research field (i.e., published journal articles, technical reports, conference proceedings etc.) have not been demonstrated.

Institutional and Temporal Confounders

There is also a special risk to the generalizability of findings within countries due to differences among the 16 Länder of Germany for example. Universities in each Länder are generally separate and have their own doctorate training programs and thus differing emphasis on disciplines and the ways in which research funds are allocated. Therefore, if a corpus of dissertation abstracts were to disproportionately sample a certain type of university (for example; Bavarian Catholic universities where there is an emphasis on family values and religious studies, or North German secular research universities where there is an emphasis on the social structure and migration) it could result in systematic bias when examining national patterns. Although the DNB corpus provides data regarding institutional distribution it is difficult to assess this confounding variable based upon the data provided.

In contrast, Turkey has a completely different but also very serious issue regarding institutional distribution. In addition to having a dramatically expanded university system since 2010 (from approximately 94 universities

in 2006 to over 200 as of 2024, including many new Anatolian universities with different research cultures and faculty compositions than the older metropolitan institutions), a corpus of dissertation abstracts using YÖK can be viewed as being comprised of very diverse institutional environments. Thus, any national-level pattern found through examining the dissertation abstracts may represent the increasing numerical predominance of the new universities (which tend to be producing more applied, practitioner-oriented research) and not necessarily a national academic culture.

Response Strategy: Institutional Stratification and Temporal Bootstrapping

Stratifying the study's reporting according to institutional type (research vs. teaching; metropolitan vs. provincial; old vs. new), as well as using institutional metadata to report institutional level findings, is the most direct way to address the within country variation issue. When institutional metadata is provided in the dataset, institutional stratification should be conducted on a routine basis and its findings should be reported along with the overall national finding(s). If institutional stratification indicates that national findings were produced primarily through findings from one institutional type, the generalizability of the findings for that nation can then be qualified based on the type of institutions producing those findings.

Temporal Bootstrapping; Resampling the corpus at regular time intervals and testing if the study's primary conclusions hold across each of the two year segments, addresses both the temporal snapshot problem and the compositional change problem (i.e., the changing composition of institutions in the Turkish corpus over time). Findings that hold for the 2019-2021 segment, the 2020-2022 segment and the 2022-2024 segment are more generally applicable than those that emerge only from an analysis of the entire six year period. Including analyses of the stability of findings in supplementary materials, as opposed to accepting the six year aggregate as definitive, raises the burden of evidence for making national level claims in cross-national bibliometrics.

Toward Epistemically Honest Cross-National Bibliometrics: A Synthesis

The Six Structural Tensions: Summary

Cross-national bibliometric analyses of doctoral research have been explored to identify and analyze six structural tensions. All six are based on systematic relationships between a methodology's structure and domains of reality that the methodology cannot be completely or wholly accessible. Table 1 summarizes the six tensions, their principal representations, and strategies for responding to these tensions.

Table 1. *Summary of Structural Tensions*

| Structural Tension | Principal Manifestation | Risk if Unaddressed | Response Strategy |
|----------------------------------|---|--|---|
| I. Metadata Dependency | Title/keyword/abstract ≠ dissertation content | Over-attribution of theoretical poverty | Full-text validation sub-sample (≥10%); stratified by country, discipline, year |
| II. Taxonomy Translation | National classification systems embed incommensurable disciplinary cuts | Spurious disciplinary contrasts as classification artefacts | Double coding; sensitivity analysis across alternative classification schemes |
| III. Visibility Asymmetry | Academic cultures differ in metadata explicitness | Systematic underrepresentation of one country's theoretical richness | Language-sensitive thesaurus expansion; second-pass classification |
| IV. Unit of Analysis | Dissertation as dated artifact; cumulative thesis format | Information loss for German cumulative theses; temporal displacement | Dissertation-publication linkage for cumulative theses; date-of-commencement coding |
| V. Policy Proximity | Co-occurrence ≠ causation; vocabulary contamination | Misleading policy-responsiveness claims | Temporal sequencing analysis; discursive context analysis on sub-sample |
| VI. Generalisability | Corpus → national academic culture generalisation | Institutional and temporal confounders misread as national patterns | Institutional stratification; temporal bootstrapping across sub-periods |

A Protocol for Epistemically Honest Cross-National Bibliometrics

The six tensions and their strategies can be reduced to a methodology protocol – a methodology standardizing reporting and design for all bibliometric comparative studies across countries. The protocol does not represent a list of required steps to take, but a methodology to provide a clear methodology for reporting on the authors' epistemology: what the study can confidently conclude, what the study can suggestively point towards, and what the study can structurally never report.

There are four parts to the methodology protocol. First, there needs to be a clearly defined Methodological Transparency Statement at the start of the Methods Section identifying the study as an "Indicator-Level" study and specifying those of the six structural tensions, that are the most relevant to the design of the study and the corpus used in the study. This will signal to the reader the appropriate interpretation register for the results (i.e., "This is how German doctoral research represents itself through metadata as collected by the Deutsche Nationalbibliothek") instead of saying "This is what doctoral students in Germany study."

Second, the Results Section should separate the findings based on their epistemic level: robust (found consistently across sensitivity analyses and validated using the validation sub-sample); provisional (consistent with expectations but influenced by classification decisions); and speculative (theoretically driven but not supported by the indicator-level data). Providing this separation will give the reader the necessary epistemological framework to evaluate which conclusions have enough evidence to be immediately translated into policy or curricular changes and which conclusions require additional empirical support before being translated into policy or curricular changes.

Third, the Discussion Section should contain a ‘Structural Limitations’ Subsection. Instead of the traditional ‘Limitations’ Paragraph placed at the end of a study as a routine acknowledgement, this Subsection would represent a critical examination of which specific Response Strategies were implemented, which Response Strategies were not (and why) and how the unimplemented Responses impact the degree to which the major conclusions of the study should be accepted.

Lastly, the Study should also contain a Public Data Appendix containing the full Coding Scheme, Thesaurus Word Lists, Translation Equivalence Tables, and Analysis Results for the Validation Sub-Sample. The purpose of providing a Public Data Appendix would be to allow both independent verification and cumulative development of the methodology – future researchers can improve upon and use the Thesaurus developed in the current study and contribute to the development of a common methodology for cross-country comparative bibliometric research, rather than needing to develop a new one for each new study.

Conclusion: The Map and the Territory Revisited

This section will explore how the “map” (a cross-national bibliometric analysis) reflects only a subset of the territory (the total set of data used to produce the analysis). While the chapter has attempted to demonstrate how the cross-national bibliometric analysis maps the territory, there are several significant ways in which the “map” differs from the “territory”. These six structural differences are not general limitations of all research. Rather they are characteristics of the indicator level epistemology that a bibliometric analysis of metadata will necessarily inhabit.

A key contribution of this chapter is not to assert that cross-national bibliometric analysis is not reliable or informative. On the contrary, I believe that it is a valuable and largely underused method for comparative education research. Cross-national bibliometric analysis is able to identify patterns in scholarly output that no other method can access at a similar scale and cost. A key contribution of this chapter is to assert that the degree to which

cross-national bibliometric analysis is valuable is directly related to the epistemological honesty with which it is practiced and presented. For example, a bibliometric study that describes the results of its indicator-level analyses as the direct result of what ‘German scholars study’, or what ‘Turkish doctoral research prioritizes’, is misleading its readers and ultimately does not serve the policy and curriculum reform debates that it wishes to support.

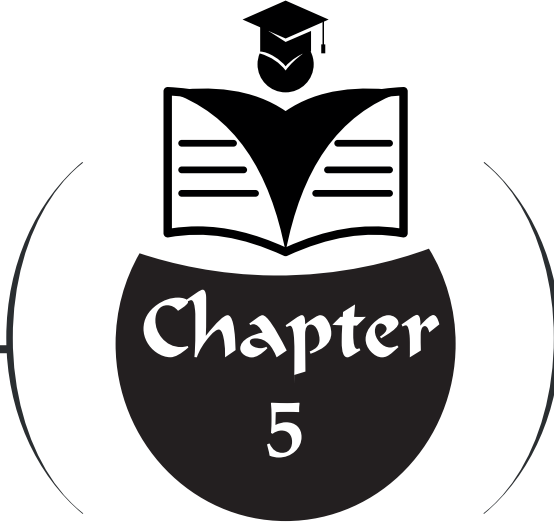
Therefore, the goal of epistemologically honest cross-national bibliometrics is to create a methodological ideal that includes mapping the shadow of the methodology alongside the findings, specifying the areas where the methodology is unable to describe what is suggested by the findings, and providing sufficient methodological transparency so that readers may interpret the findings accordingly. The strategies for responding to these limitations that have been provided in this chapter provide a single path toward this ideal. A second goal is to develop a comparative education bibliometrics community that views methodological reflection as integral to empirical work rather than simply as an appendix to empirical work. The map, accurately and truthfully drawn, remains a potentially powerful navigational tool, even when, and possibly particularly when, it clearly indicates the uncharted waters that exist beyond its boundaries.

REFERENCES

- Arabacıoğlu, B. (2020). Edebi çeviri ve çeviri eleştirisi: Karşılaştırmalı bir çalışma [Literary translation and translation criticism: A comparative study]. *Anadolu Journal of Educational Sciences International*, 10(2), 1004–1016. <https://doi.org/10.18039/ajesi.724391>
- Arabacıoğlu, B., & Balkaya, Ş. (2023). Literatur und interkulturelle Sensibilisierung: Ein psychotherapeutischer Roman im Hinblick auf Vorurteile, Stereotype und Fremdbilder in einer multikulturellen Gesellschaft. *International Journal of Language Academy*, 11(3), 297–315. <http://dx.doi.org/10.29228/ijla.71805>
- Arabacıoğlu, B., & Tunçyüz, S. N. (2025). Künstliche Intelligenz in der Fremdsprachendidaktik: Sprachgefühl und Sprachmodell im Fokus. *Diyalog, Sonderausgabe: Germanistik im 21. Jahrhundert – Band I*, 413–429. <https://doi.org/10.37583/diyalog.1824338>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Başaran, B. (2025). Ailenin eğitimdeki rolü üzerine akademik temsiller: Türkiye ve Almanya doktora tezlerinde karşılaştırmalı bir inceleme (SÇB-2025-2934 Sonuç Raporu). *Anadolu Üniversitesi BAP Koordinasyon Birimi*.
- Bernstein, B. (1975). *Class, codes and control: Vol. 3. Towards a theory of educational transmissions*. Routledge & Kegan Paul.
- Bourdieu, P. (1988). *Homo academicus*. Polity Press.
- Bray, M., & Thomas, R. M. (1995). Levels of comparison in educational studies: Different insights from different literatures and the value of multilevel analyses. *Harvard Educational Review*, 65(3), 472–490. <https://doi.org/10.17763/haer.65.3.g3228437224v4877>
- Bray, M., Adamson, B., & Mason, M. (Eds.). (2014). *Comparative education research: Approaches and methods* (2nd ed.). Springer.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Çırak, A. (2022). Sosyal bilgiler ve karşılaştırmalı eğitim: Türkiye’de sosyal bilgiler eğitimi alanında farklı ülkelerle karşılaştırmalı olarak yapılan tezlerin incelenmesi. *Eğitim ve Toplum Araştırmaları Dergisi*, 9(1), 208–228. <https://doi.org/10.51725/etad.1087300>
- Dolowitz, D. P., & Marsh, D. (2000). Learning from abroad: The role of policy transfer in contemporary policy-making. *Governance*, 13(1), 5–23. <https://doi.org/10.1111/0952-1895.00121>
- Glanzel, W., & Schoepflin, U. (1999). A bibliometric study on ageing and reception processes of scientific literature. *Journal of Information Science*, 21(1), 37–53. <https://doi.org/10.1177/016555159502100104>

- Kayhan Yüksel, D., & Baka, B. (2022). Thematic and methodological trends of graduate thesis in comparative education: A content analysis. *Journal of Computer and Education Research*, 10(20), 777–794. <https://doi.org/10.18009/jcer.1174500>
- Korzybski, A. (1933). *Science and sanity: An introduction to non-Aristotelian systems and general semantics*. Institute of General Semantics.
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLOS ONE*, 10(6), Article e0127502. <https://doi.org/10.1371/journal.pone.0127502>
- Lietz, P. (2010). Research into questionnaire design: A summary of the literature. *International Journal of Market Research*, 52(2), 249–272. <https://doi.org/10.2501/S147078530920120X>
- Marginson, S. (2016). High participation systems of higher education. *Journal of Higher Education*, 87(2), 243–271. <https://doi.org/10.1353/jhe.2016.0007>
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Noyons, E., Moed, H. F., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study. *Journal of the American Society for Information Science*, 50(2), 115–131.
- Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25(4), 348–349.
- Sarkiler, Y. A. (2025a). CEFR alignment in German textbooks: A comparative analysis of publisher claims, expert opinions, and AI evaluations. In M. Özüdoğru (Ed.), *Innovative applications and research methods in education sciences* (pp. 166–182). Platanus Publishing. <https://doi.org/10.5281/zenodo.18056942>
- Sarkiler, Y. A. (2025b). Evaluating DaF textbooks: Structure, teaching practice and CEFR. In S. Demir Başaran (Ed.), *Theoretical and applied approaches in educational sciences* (pp. 24–36). Duvar Yayınları. ISBN: 978-625-5885-40-1
- Sarkiler, Y. A. (2025c). Sustainability in foreign language learning in Turkey: Approaches, challenges and perspectives for a future-oriented language education. In F. Çelebi (Ed.), *Theory, context, and practice in social, human, and administrative sciences* (pp. 61–73). Duvar Yayınları. ISBN: 978-625-5698-22-3
- Tatlı, S., & Adıgüzel, O. C. (2012). Türkiye'deki lisansüstü karşılaştırmalı eğitim tezlerinin çok boyutlu bir incelemesi. *Anadolu Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12(1), 143–150.
- Tight, M. (2019). *Documentary research in the social sciences*. SAGE.
- van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62(1), 117–131. <https://doi.org/10.1007/s11192-005-0007-7>

- Whitley, R. (2000). *The intellectual and social organization of the sciences* (2nd ed.). Oxford University Press.
- Yılmaz Gngr, Z. (2025). Yabancı dil ğretiminde teknolojik aralar. In Z. Yılmaz Gngr (Ed.), *Fransız dili eđitiminde alıřmalar* (pp. 25–45). BİDGE Yayınları. ISBN: 978-625-372-672-0



REFLECTIONS OF ALTERNATIVE EDUCATION EXPERIENCES ON STUDENTS' ACADEMIC PERFORMANCE AND SKILL DEVELOPMENT¹



Ece YILMAZ²

Pınar ÇAVAŞ³

Aşlı Ceren ALAÇAM⁴

1 This chapter is derived from the master's thesis titled Başka Bir Okul Mümkün (BBOM) Eğitim Modeline Göre İlkokul Eğitimi Almış Öğrencilerin Ortaokul Adaptasyonlarının İncelenmesi by Ece Yılmaz (Ege University, Institute of Educational Sciences, 2024), supervised by Prof. Dr. Pınar Çavaş.

2 Primary School Teacher- Özel Renkli Orman İlkokulu, İzmir, Türkiye
ecee9597@gmail.com, ORCID: 0009-0002-4441-6676

3 Prof. Dr., Ege University, Faculty of Education, Department of Primary Education, İzmir, Türkiye, pinar.cavas@ege.edu.tr, ORCID: 0000-0001-9492-9002

4 Dr., Research Assistant, Ege University, Faculty of Education, Department of Primary Education, İzmir, Türkiye, ORCID: 0000-0003-4152-0902

1. INTRODUCTION

Education is a living system that ensures social continuity and evolves simultaneously with the socio-cultural dynamics in which it is embedded. With the construction of modern nation-states, education has assumed a mission beyond merely preparing individuals for the future; it has become a vehicle for maintaining social standards and fostering “ideal citizens” who have internalized national ideals. In this process, central authorities have institutionalized the “mainstream” educational model by standardizing all components, from curriculum design and teacher competencies to institutional operations and assessment mechanisms. However, the homogeneous structure offered by mainstream education today brings about significant pedagogical and ethical debates. Standardized educational models tend to perceive individual differences as “deviations” to be managed rather than assets to be celebrated, often labeling learners within predefined molds. This approach, rooted in classical behaviorist theory, views learning as a linear process based on stimulus-response bonds that is expected to occur uniformly for everyone (Kaya & Gündüz, 2015). This mechanistic perspective may result in the stifling of creativity, the restriction of critical thinking skills, and, most importantly, the reproduction of social inequalities through education.

As an antithesis to the limitations created by the traditional educational paradigm, alternative education models redefine the learning process on a completely different ontological ground. These models offer a structure that prioritizes flexibility, student autonomy, and democratic participation instead of hierarchical and rigid programs. Approaches such as Montessori, Waldorf, Reggio Emilia, and its unique reflection in Turkey, BBOM (Another School is Possible), place individual differences at the heart of pedagogy. They promise an “ecological” educational climate where every child can develop at their own pace and in their own style. In this context, alternative education represents not merely a change in methodology, but a radical shift in mindset regarding human nature, the right to learn, and social peace.

A growing body of research on alternative education models indicates that these pedagogical approaches can have significant positive effects on students’ cognitive, social, and emotional development. Emerging as alternatives to the standardized and often highly structured nature of mainstream educational systems, alternative education models prioritize student autonomy, experiential learning, and learner-centered pedagogical practices (Raywid, 1994; Sliwka, 2008). Within such educational environments, students are encouraged to actively construct knowledge through inquiry, collaboration, and self-directed learning rather than through passive reception of information. These models typically emphasize individualized learning pathways and holistic development, fostering both academic competence and personal growth.

Empirical studies have demonstrated that students educated within alternative pedagogical frameworks frequently exhibit strong academic and socio-emotional outcomes. For instance, research conducted by Lillard (2005) shows that students who attend Montessori schools tend to outperform their peers in traditional educational settings in areas such as mathematical reasoning, literacy skills, and executive functioning. Similarly, other studies have highlighted that alternative education environments often support higher levels of student engagement, intrinsic motivation, and self-regulated learning (Rathunde & Csikszentmihalyi, 2005). These outcomes are largely attributed to the emphasis placed on learner autonomy, collaborative classroom cultures, and democratic participation in educational processes.

Beyond academic achievement, alternative education practices implemented at the primary school level are also recognized for their contribution to the development of competencies commonly associated with twenty-first century skills. Critical thinking, creativity, collaboration, communication, and complex problem-solving abilities are frequently nurtured within these learning environments. Such competencies are supported through inquiry-based learning practices, interdisciplinary curricula, and classroom environments that allow students to take greater responsibility for their own learning processes (Biesta, 2015; Kraftl, 2013). The democratic structure of many alternative education settings further promotes students' sense of agency and participation, which in turn contributes to their social and emotional development.

Despite these documented advantages, the transition of students from alternative primary education environments into the more structured and hierarchical contexts of mainstream secondary schools remains a relatively complex and underexplored issue in the literature. Secondary school environments are typically characterized by standardized curricula, subject-based teaching structures, increased academic expectations, and more formal evaluation systems. As a result, students who previously experienced flexible and learner-centered educational environments may initially face challenges in adapting to these institutional expectations. Miller (2007) suggests that students transitioning from alternative education contexts may experience an initial "adaptation shock," primarily resulting from differences in educational philosophy, classroom interaction patterns, and assessment practices. However, the same research also indicates that students educated in alternative learning environments often possess higher levels of intrinsic motivation, adaptability, and self-regulation, which may help them overcome these initial challenges and achieve positive long-term academic and social outcomes.

The dynamics of such educational transitions can be further understood through the **Stage-Environment Fit Theory** proposed by Eccles and

colleagues (Eccles et al., 1993). According to this theoretical framework, students' motivation, engagement, and well-being are strongly influenced by the degree of alignment between their developmental needs and the characteristics of their learning environments. When there is a strong fit between students' developmental stage and the opportunities provided by the educational context, students are more likely to experience positive academic and socio-emotional outcomes. Conversely, when there is a mismatch between these factors, students may encounter difficulties in adapting to the educational environment. Educational transitions, such as the shift from primary to secondary school, may therefore create situations in which the previously established alignment between students' learning experiences and the new institutional environment becomes disrupted.

Within the Turkish educational context, interest in alternative education models has increased in recent years as part of broader discussions on educational innovation, diversity, and child-centered pedagogy. Initiatives such as the **Another School is Possible (Başka Bir Okul Mümkün – BBOM)** movement have aimed to create educational environments that emphasize democratic participation, ecological awareness, and student-centered learning. Despite the increasing visibility of such initiatives, empirical research examining the long-term educational experiences of students educated within these models remains limited in Türkiye. In particular, there is a notable lack of studies investigating how students graduating from alternative primary schools adapt to conventional secondary schools characterized by exam-oriented curricula and more structured institutional expectations.

The Purpose of the Study

The purpose of this study is to examine the academic performance and skill development of students who graduated from İzmir Private Renkli Orman Primary School—an institution implementing the BBOM (Another School is Possible) educational model as an alternative to mainstream schooling—following their transition to middle school. In line with this purpose, the study seeks to answer the following research questions:

- How is the academic performance of students who received alternative education at the primary level after transitioning to mainstream schools?
- Which specific skills have been developed by students who received alternative education during their primary school years?

2. METHOD

Research Design

This study was conducted using the case study design, one of the qualitative research models, as it aims to obtain descriptive and explanatory

information regarding alternative education by consulting the perspectives of graduates from İzmir Private Renkli Orman Primary School (which adopts the Another School is Possible – BBOM model), their parents, and their middle school teachers.

The case study is a qualitative research design that involves an in-depth investigation of a single case, individual, event, setting, or document. In this method, researchers employ various data collection and analysis techniques to understand the complexity of a specific phenomenon. By focusing on a single case, case studies provide an intensive analysis through qualitative data and seek to understand how the context influences the situation by examining it within its natural environment (Yıldırım & Şimşek, 2018).

Participants

The purposive sampling method was employed to determine the study group of the research. Purposive sampling is a technique that allows researchers to select a sample based on specific criteria. Unlike random sampling, this method enables the selection of individuals or cases that possess predefined characteristics relevant to the study. Thus, the researcher can obtain in-depth and comprehensive information through a smaller number of participants (Yıldırım & Şimşek, 2018). For this purpose, the study group consisted of 5 graduates from İzmir Private Renkli Orman Primary School (S1, S2, S3, S4, S5), 6 parents (P1, P2, P3, P4, P5, P6) and 7 subject teachers (T1, T2, T3, T4, T5, T6, T7) currently instructing these students in middle school.

In the selection of the study group, only students attending the same middle school were included. Although the researcher intended to interview an additional branch teacher from the middle school, the study remained limited to the aforementioned participants as that teacher declined to participate.

Instrument Development

Semi-structured interview protocols were developed to examine the adaptation processes of students who had received education through an alternative education model in primary school. The questions were structured around three main themes derived from the research questions: academic, social, and emotional adaptation. Separate interview forms were prepared for students, parents, and subject teachers to obtain perspectives from different stakeholders in the adaptation process. To ensure content validity and clarity, expert opinions were obtained from a primary school teacher experienced in alternative education and a lecturer specializing in curriculum development.

A pilot interview was conducted with the parent of a student who had previously studied in the same primary school classroom but had transferred to another school in the fourth grade. Based on the pilot study, several

questions were revised and probing questions were added to improve clarity and depth.

The final interview protocols consisted of 16 questions for subject teachers, 12 questions for students, and 7 questions for parents. In addition, subject teachers were provided with a brief explanation of the alternative education model implemented at Private Renkli Orman Primary School to ensure a clear understanding of the research context.

Data Collection

Data were collected through semi-structured interviews with students, parents, and subject teachers. Individual interviews were conducted with teachers and parents, while focus group interviews were conducted with students. Participation was voluntary, and participants were informed about the purpose of the study and the confidentiality of their responses. All interviews were audio-recorded with participants' consent. The recordings were transcribed verbatim by the researcher using Microsoft Word and used as the primary data source for analysis. Throughout the process, efforts were made to maintain researcher neutrality and minimize potential bias.

Data Analysis

The data collected through semi-structured interviews were analyzed using descriptive analysis, guided by the core research questions. The primary objective of this approach is to present the findings to the reader in an organized and interpreted manner. The interview questions were structured around themes established by the researcher following an extensive literature review in alignment with the research problem. Data analysis was conducted based on these predefined themes. In the descriptive analysis process, direct quotations were frequently incorporated to effectively convey the participants' perspectives (Yıldırım & Şimşek, 2018). The data were categorized, summarized, and interpreted according to the established thematic framework. Significant views from certain participants were presented verbatim to maintain their original meaning.

Trustworthiness

In qualitative research, validity and reliability are addressed through various verification strategies (Yıldırım & Şimşek, 2018). In this study, data triangulation was employed to enhance the credibility of the research. Data obtained from students, parents, and teachers were evaluated together to obtain a more comprehensive understanding of the phenomenon.

The semi-structured interview forms used in the study were developed based on the opinions of three academics specialized in qualitative research and one primary school teacher experienced in alternative education. Following

pilot implementation, the interview forms were revised and finalized.

In order to preserve the authenticity of participants' perspectives during the data analysis process, direct quotations were included in the findings. All interviews were audio-recorded and transcribed verbatim. The collected data were securely stored throughout the research process to ensure confirmability.

Ethical Considerations

Before the data collection process began, ethical approval was obtained from the relevant ethics committee, and official permission to conduct the research was granted by the İzmir Provincial Directorate of National Education.

All participants, including students, parents, and teachers, were informed about the purpose and scope of the study, and participation was based on voluntary consent. Parental consent was obtained for student participants.

Throughout the research process, the confidentiality of participants' identities was carefully protected, and the collected data were anonymized during reporting. All stages of the study were conducted in accordance with the principles of scientific research and publication ethics.

3. FINDINGS

This section presents the findings obtained from the analysis of the interview data conducted to answer the research questions. In line with the purpose of the study, the findings were examined under two main categories: "views on students' academic performance" and "views on the development of students' cognitive and social skills" during their transition to middle school after receiving alternative primary education.

The data obtained through semi-structured interviews were analyzed using descriptive analysis, and participants' views were organized under specific themes. The findings are presented and interpreted with the support of direct quotations from participants.

3.1. Views on Students' Academic Performance

To address the research question "How is the academic performance of students who received alternative education at the primary level after transitioning to mainstream schools?", interviews with students, teachers, and parents were analyzed. The findings indicated that participants' views were organized around three main themes: readiness, academic achievement, and assessment and evaluation processes.

Readiness

Participants' views indicate that although the students came from an alternative education background, their academic readiness when they started

middle school was sufficient and, in some cases, above the class average. Statements from parent P4 and teachers T3 and T6 support this finding.

P4: When she transitioned from an alternative school like Renkli Orman to a public school, she was already familiar with everything in the curriculum, actually even more. There was never a moment in any course where she said, “We were supposed to learn this, but we never covered it.” She never fell below the class average in any subject. Even though Renkli Orman is an alternative school where there is generally less homework, almost none at all, and she moved into a system where homework is common, she never had any difficulty taking responsibility for those assignments.

T3: They already knew everything I taught. In fact, they would say that they had already seen these topics last year.

T6: There were no topics that they lacked knowledge about.

These views indicate that when the students started middle school, they were generally well prepared for the topics included in the curriculum.

In the views regarding readiness levels, differences were more evident in certain subjects. In particular, it was stated that students coming from Renkli Orman were at a more advanced level in English and Social Studies. Some participants expressed this situation as follows:

For example, S3 described the English level in the current school by saying, “English here is a little easier.” Similarly, S5 stated, “We have just started covering topics in English that we had already learned before,” indicating that they were ahead in this subject. Teacher T7 commented in a way that supported the students’ views: “I was quite satisfied with their English level... the students had high self-confidence and a high rate of participation in class. Their speaking and practical skills were also very good.”

These views suggest that students coming from Renkli Orman had a higher level of readiness particularly in English. This may be explained by the fact that Renkli Orman uses its own English curriculum.

Similarly, it was also stated that students were ahead of their classmates in certain Social Studies topics. Student S5 expressed this as follows: “Academically, we were definitely a bit ahead in Social Studies at Renkli Orman. Because when we were in fourth grade, we had already learned the systems of government that they have only covered partially here so far. We had learned them in a much more detailed way there.” Teacher T2 also commented on this subject, referring to all of the students from the alternative school: “All five of them are above the class average.”

However, some teachers also noted that there were individual differences among students. Teacher T1 expressed this by saying, “There were some

students who were academically behind, with difficulties in writing and reading. We needed to work with them more closely.” This statement indicates that although there were individual differences among students, the areas identified as weak were gradually improved with support over time.

From the students’ perspective, certain subjects were mentioned in which they felt academically less competent. These subjects included Folk Culture, Music, and Visual Arts.

Student S2 stated, “Sometimes the essays we write in the Folk Culture course are not really related to the topic. When we read them again later, they do not make much sense.” Similarly, S1 said, “I think they do not really value music or visual arts here. Neither the students care much about it nor is there a teacher available to teach it.”

These statements suggest that students perceive differences between the learning experiences they had in the alternative education environment and the instructional practices in their current school.

Academic achievement

When the participants’ views were examined, it was observed that students’ academic achievement levels were generally evaluated as high. Both students and parents stated that students coming from the alternative education environment typically performed above the class average in their classrooms. Student S4 expressed this situation as follows: “...In this school there are practice exams, and in those exams the students who usually achieve the highest scores are the ones who graduated from Renkli Orman.” Parent P3 also commented on their child’s performance: “...In the nationwide exams they also give rankings at the provincial and national levels. My child performs well there as well. He or she is usually first or second in the class and ranks around 150–200 across İzmir.”

Similarly, Parent P4 pointed to their child’s success by stating: “When she transitioned from an alternative school like Renkli Orman to a public school, she was already familiar with everything in the curriculum—even more than required. She never fell below the class average in any subject.”

These views suggest that students coming from the alternative education environment generally perform above the class average in terms of academic achievement.

Assessment and evaluation processes

Another theme that emerged from the interviews regarding students’ academic performance was related to assessment and evaluation processes. Within this theme, two categories were identified: examinations and homeworks.

Examinations

When the teachers' views were examined, it was observed that the exam performance of students coming from Renkli Orman was generally evaluated positively. Some of the teachers' views regarding this issue are presented below.

T1: "They raised the class average in the exams. Some students even scored higher than the top two students in the class. Therefore, even though they did not come with exam anxiety and had not previously experienced exam pressure or obligations, they entered this competitive environment quite easily. They did not have any problems with grades; they showed responsibility in getting good grades, studying, and fulfilling their duties through their own efforts. I never felt the need to meet separately with their parents; they already had this level of maturity when they arrived."

T7: "When I compare them with the other students, almost all of them were at a very good level."

T4: "They are at the same level as the rest of the class. One student might be slightly ahead, but overall they are well aligned with the class."

These statements indicate that teachers generally evaluate the students' exam performance positively.

During the interviews, teachers also stated that they did not observe a strong level of exam anxiety among the students, and that even when some anxiety was present at the beginning, it disappeared over time. Some of the teachers' views regarding this issue are presented below.

T7: "...they did not really have much exam anxiety. At first they did not know much about exams, so they felt a little nervous and worried about grades and similar things. But after the first exam, once they understood what each teacher expected and how the procedure worked, and since they did not have any knowledge gaps, they adapted very easily. There was no problem."

T3: "Since I personally did not want them to feel anxious, I did not observe such anxiety among them."

T2: "They adapted to the exam system just like the other students. If there is an exam, they come prepared and study for it. They do not have difficulty answering exam questions, solving tests, or responding to other question types such as true-false, matching, fill-in-the-blank, or open-ended questions. I did not observe any difficulties in these areas."

T1: "At the beginning, there were a few students who were a bit surprised by the process, but after individual meetings with them and their parents, they overcame these difficulties. Now they do not have any noticeable differences or fears regarding achievement."

T4: “At first they were anxious about the Religious Culture course. They were asking questions such as ‘How will you ask the questions?’ and ‘What kind of questions will there be?’ But now that anxiety is gone. They have become accustomed to the system.”

In alternative education environments, assessment is typically conducted through methods other than exam-based evaluation. Nevertheless, the teachers’ views suggest that students adapted to the exam-based assessment processes in middle school within a relatively short period of time.

Homeworks

Another dimension of the assessment and evaluation process concerns homework. In the views related to homework, parents stated that students did not experience difficulty in taking responsibility for their assignments. Some parental views regarding this issue are presented below.

P4: “In Renkli Orman there is very little homework—almost none at all. Despite this, when my child transitioned into a system where homework is common, she did not experience any difficulty in taking responsibility.”

P3: “My child has a highly developed sense of responsibility. This may be partly genetic and partly related to the BBOM culture.”

P5: “She does her homework. We do not experience many difficulties; she manages and follows her responsibilities on her own without much struggle.”

These parental views were also supported by the statements of the teacher participants.

T2: “If I evaluate them in general, they are again above the class average in terms of responsibility. They fulfill their homework and other duties.”

T1: “There is a certain awareness among these students. They do not have the feeling of ‘I cannot do it’; instead, they say ‘I can do it.’

T4: “They complete their homework... There are some classmates who bring incomplete work, but these students bring all their materials and complete their assignments on time. For example, during a school trip, two students even submitted their homework on the way in order to fulfill their responsibilities.”

Another noteworthy finding related to homework is that students appear to develop different perspectives when completing their assignments. Some teachers’ views on this issue are presented below.

T1: “These students have a different perspective. When preparing projects, they put their own effort into the work. They never do simple copy–paste work or just take something from the computer.”

T3: “S5 and S2 are students who think somewhat differently; they can solve problems in different ways.”

T4: “In terms of homework, I particularly admire the creativity of one student. This student adds something of his or her own. Of course, other students can also contribute their own ideas, but this student has a unique approach. This student also has a strong ability in drawing and visual expression. Another student is similar in a different way; there is a calmness in how the work is completed. You can clearly see the order and care in the way the homework is done.”

In addition to these views, there are also statements indicating that students are particularly willing to engage in tasks requiring creative thinking and manual skills.

T5: “...when the task requires handicraft work, they are usually ahead of the others.”

T6: “They are very willing in this regard (manual skills). For example, last year I asked them to create landforms using clay. They all brought their materials, prepared the cardboard, and colored them as they wished, of course in accordance with the correct colors of the landforms. We created maps in this way and even wrote cuneiform on clay tablets. They enjoyed it and completed the work willingly and enthusiastically.”

3.2. Views on Students’ Cognitive and Social Skills

In order to answer the research question, “Which skills have been developed by students who received alternative education during their primary school years?”, the interviews conducted with teachers were analyzed. The analysis revealed that the views were grouped under three main themes: problem-solving skills, thinking skills, and communication skills. Examples of views related to each theme and their interpretations are presented under the relevant headings.

Problem-Solving Skills

During the interviews with teachers, when they were asked how they evaluated students’ problem-solving processes, it was observed that students tended to try to resolve problems among themselves first. Some of the participants’ views are presented below.

T2: “...In minor conflicts, while other students usually complain about each other or give names, these students tend to support one another. Problems do not occur very often, and when they do, they do not immediately report each other’s names... Instead, they come together and talk among themselves... These students are more conscious. They are more aware of events and situations; they are more mature. They first think things through

and evaluate the situation in their minds and then take action. There is no sudden reaction or acting without thinking.”

T3: “...A classmate bit him and it caused bruising. He kept showing it to me and saying that it hurt, but he did not complain about his friend. The next day I saw them together again; they had managed to solve the issue on their own.”

These statements indicate that students tend to try to resolve the problems they experience among themselves first and show a tendency to handle conflicts without immediately escalating them to others. It can also be understood that the culture of resolving problems collaboratively among peers was carried over from their previous school into the middle school environment.

The interviews also included statements indicating that when students were unable to resolve problems among themselves, they sought support from their classroom teachers or the school administration. Some of the participants’ views on this issue are presented below.

T1 attributed this attitude of the students to the fact that their previous school was accustomed to this type of problem-solving: “Whenever a problem occurs, they first try to solve it among themselves. For example, if they have a conflict with a friend, they first try to warn each other and solve it on their own. If they cannot resolve it, they come and tell me about it in a very calm and respectful way... Sometimes they even ask another friend to mediate and help solve the problem. In most cases they resolve their problems themselves. Because they are already accustomed to handling issues this way, most problems do not reach us and fade away before that stage. Of course, if they cannot overcome the problem, then they come to us and we try to help as much as we can.”

Similarly, T6 stated: “Of course they try to resolve issues among themselves first, but when they cannot, they become upset. Because they are upset, the issue eventually comes to us and turns into a complaint. Then we try to solve it together. When they cannot resolve it among themselves, we solve it together; otherwise it does not work. There was an incident last year when a student from Renkli Orman became very upset and even cried because he felt he had been treated unfairly. But we solved the issue in the end. They communicated with each other, reconciled, and continued getting along.” T6, with this expression, describes children’s problem-solving processes and how they turn to adults when they are stuck.

Communication Skills

Another theme that emerged from the interviews with teachers was the development of students’ communication skills. Some teacher views reflecting this evaluation are presented below.

T1 stated: “First of all, their communication skills are at a high level. I think this may also be an advantage provided by the system they previously experienced. You can see their politeness and gentleness in their relationships with their friends... When an environmental or social issue is discussed in class, they always express their sensitivity about it... The way they express themselves, their respect for teachers, their respect for friends, and their respect for the environment and all living beings is remarkable... They have contributed a lot to us in terms of communication skills.” Through this statement, the teacher listed many positive aspects of the students’ communication behaviors and suggested that these characteristics might be related to the education they previously received.

Another teacher, T5, emphasized students’ ability to express themselves without resorting to complaints: “They can explain themselves very well in class and when discussing a topic. When they have a problem, they can express it clearly. Other students usually complain about everything, but these students do not complain much. If there is a real problem, they come and explain it.”

Similarly, T2 stated, “Their way of expression is very clear; they speak openly,” while T4 commented, “They can express themselves directly and clearly.” These statements indicate the students’ competence in expressing themselves.

Participant T3 also described the advanced level of these students’ communication skills: “...They are not like typical children... They seem more mature for their age; sometimes they can understand you even without you saying anything, sometimes just through eye contact... I think they express themselves more comfortably and with more appropriate words. They can state their opinions without hesitation and do so without hurting others.”

Based on the teachers’ statements, it can be understood that students coming from Renkli Orman are able to express themselves openly, comfortably, and politely, and that they demonstrate sensitivity and self-confidence in their interactions.

At the same time, the interviews also revealed that there were individual differences in the communication skills of students who graduated from Renkli Orman. The views of participants T7 and T6 regarding this issue are presented below.

T7: “S4 used to be somewhat introverted before. She spoke more quietly and calmly. Now she is better. S2 was similar, but now they are both doing better. They speak more confidently now, and the shyness in their voices has mostly disappeared.”

T6: “Some of our students can express themselves very clearly. I think it also depends on personality. According to this educational system, their expression skills seem to have improved, but while some students express themselves very openly and clearly, one or two students are more reserved and have difficulty expressing themselves... There are still some shy students. They are not all completely extroverted. I think this is related to individual differences.”

Thinking Skills

In the evaluations of students’ cognitive skills expressed during the interviews with teachers, one of the prominent themes was that students possess advanced thinking skills. The views of participants T1, T4, T6, and T7 presented below were evaluated within the scope of this theme.

T1 stated: “These students stimulate writing, communication, and thinking skills in the classroom environment. They also encourage the ability to produce something new. If there is a point where they feel they are lacking, they try to recognize it and work to improve it. All of these are remarkable advantages. Of course, family influence may also play a role here, as well as the educational system.”

T4 commented: “...For example, they are currently writing books. Using their imagination, they create their own drawings and even invent their own names.”

T6 stated: “Their thinking skills are stronger; they come with more developed questioning skills and imagination, and they have different perspectives. They have their own imaginative worlds. They create books and support them with cartoons; they even create their own planets and give them names... Other students in the class have also started to adapt to this. As they continued doing these things, other students began writing books and creating their own imaginary worlds as well. They became a positive example within the classroom.”

T7 stated: “...They are more creative and more colorful in their thinking. They can combine different elements and produce more analytical and well-developed solutions.”

Different teachers pointed out that students demonstrate advanced abilities in creative thinking, critical and questioning thinking, multidimensional thinking, and analytical reasoning. Based on the teachers’ statements, it can be understood that students coming from Renkli Orman display well-developed thinking skills. Teachers also indicated that these students serve as positive role models for other students in the classroom.

4. DISCUSSION AND CONCLUSION

In this study, the academic performance and skill development of students who received their primary education within an alternative education model were examined following their transition to middle school. The findings indicate that the experience of alternative education has significant implications for students' academic adjustment processes as well as their social and cognitive skills.

Students' Academic Performance

The findings of the study indicate that students who received their primary education within an alternative education model demonstrate adequate academic readiness when transitioning to middle school and, in some subjects, even perform above the class average. This finding contradicts the common assumption that students coming from alternative schools, where educational practices differ from those in mainstream public schools, may experience difficulties in adapting to middle school environments.

In addition, the findings suggest that students who were educated at Renkli Orman (RO) are more willing and motivated to participate in classroom activities and learning tasks. This may be associated with the student-centered learning approaches implemented in alternative educational environments, which encourage students' active engagement in the learning process.

Teachers, students, and parents generally reported that students educated at RO perform either above the class average or close to it in terms of overall academic achievement. Participant views particularly highlighted that these students demonstrate more advanced levels in subjects such as English and Social Studies. The fact that Renkli Orman implements its own English curriculum, as well as the greater opportunities for experiential learning in Social Studies commonly found in alternative schools, may help explain the relatively stronger performance observed in these areas.

These findings suggest that the alternative education model may positively influence students' academic achievement and emphasize the importance of expanding such models and enabling teachers to use these approaches more effectively. Research in the literature also indicates that alternative teaching methods can significantly improve students' academic performance compared to traditional instructional methods. A meta-analysis conducted by Tutal, Kaçire, and Atabey (2016) found that approaches such as educational games, inquiry-based learning, project-based learning, discussion methods, and case studies are among the most effective instructional strategies. The results of the present study support the view that such practices, commonly used in alternative education settings, may play an important role in teacher education and contribute to improved student achievement.

The findings also reveal that individual differences exist among students. Some teachers indicated that certain students required additional support, particularly in reading and writing skills. However, these difficulties were reported to have been largely addressed through teacher guidance and students' adaptation processes. This suggests that even when students from alternative educational environments experience certain challenges, they are able to employ compensatory skills that help them overcome these difficulties.

Students themselves also expressed feelings of inadequacy regarding their academic performance in certain subjects, particularly Folk Culture, Music, and Visual Arts. However, it was observed that students tended to attribute these difficulties not to their own abilities but rather to the perceived inefficiency of the lessons. This indicates that these challenges are not necessarily related to disadvantages stemming from the educational environment in which the students previously studied.

From the perspective of assessment and evaluation processes, the findings suggest that students were able to adapt relatively quickly to the exam-based assessment system in middle school. Teachers' statements indicate that students generally performed well in exams and did not experience significant levels of exam anxiety. The alternative education model may positively influence students' attitudes toward assessment processes and homework responsibilities. Students' ability to participate in evaluation processes without experiencing exam anxiety may be associated with the autonomy and intrinsic motivation supported in alternative educational environments. According to Self-Determination Theory, students who have greater control over their learning processes tend to experience lower academic anxiety and higher learning motivation (Ryan & Deci, 2000). Participants' views regarding students' exam performance and their relationship with exam anxiety suggest that practices aimed at reducing exam anxiety in RO may have been effective.

The findings related to homework processes also indicate that students demonstrate well-developed responsibility-taking skills. Both parents and teachers stated that students regularly complete their homework and are able to display original approaches in the process of producing assignments. This situation may be interpreted as a positive outcome of alternative educational environments that support students' responsibility-taking and self-regulation skills.

Development of Cognitive and Social Skills

The findings of the study indicate that students who received education in an alternative learning environment also demonstrate advanced characteristics in terms of cognitive and social skills.

Teachers' views reveal that students behave more consciously in their

problem-solving processes. Their tendency to first attempt to resolve problems among themselves, and to seek support from teachers or school administrators when they are unable to do so, suggests that these students display responsible and independent behavior. This situation may be interpreted as being related to the development of responsibility-taking and collaboration skills among students educated in alternative learning environments. Indeed, research shows that problem-based learning processes in student-centered learning environments contribute to the development of independent problem-solving and collaborative skills (Hmelo-Silver, 2004).

When evaluated in terms of communication skills, the majority of teachers stated that students are able to express themselves openly and comfortably while using a respectful and sensitive communication style. These findings are also supported by literature. For instance, Johnson and Johnson (2009) emphasize that alternative education programs strengthen students' communication and social skills. Similarly, the findings are consistent with the research of Durlak, Weissberg, Dymnicki, Taylor, and Schellinger (2011), which demonstrates that social-emotional learning programs enhance students' academic and social competencies. The findings suggest that students educated in alternative learning environments are able to participate actively in group work and integrate successfully with their peers after transitioning to mainstream schools. In this respect, the results can be interpreted as evidence of the contributions of alternative education programs, which emphasize collaborative learning and group work, to the development of students' social communication skills.

In addition, teachers' views indicate that students demonstrate well-developed thinking skills. Participants highlighted students' abilities in creative thinking, imagination, questioning, and developing multiple perspectives. Their creativity and ability to use different strategies in problem-solving can also be considered indicators of advanced thinking skills. These findings may be interpreted as a result of the learning experiences provided in alternative education environments that support students' critical and creative thinking. Some teachers also stated that these characteristics enable such students to serve as positive role models for their classmates. This suggests that the experiences gained through an alternative education model may have the potential to create transformative effects for students in their subsequent educational environments.

The findings of the study can also be interpreted within the framework of Stage-Environment Fit Theory (Eccles et al., 1993) in terms of understanding how students educated in alternative environments adapt to middle school settings. According to this theory, when there is a strong alignment between students' developmental needs and the characteristics of the learning environment, students' academic engagement, motivation, and social

adjustment tend to improve. In this respect, it can be argued that previous learning experiences may play a significant role in students' transitions between different educational environments. In the present study, the fact that students who received education in alternative learning environments demonstrated well-developed problem-solving, communication, and thinking skills and were able to adapt quickly to their new schools may be interpreted as evidence that their prior educational experiences provided a learning environment that supported active participation in learning processes.

Overall, the findings of the study provide strong evidence that students who received their primary education within an alternative education model are able to adapt to middle school environments both academically and socially. The fact that students demonstrate adequate academic readiness, adapt successfully to exam and homework processes, and develop a strong sense of responsibility suggests that the learning experiences gained in alternative education environments may have positive effects on later stages of education. These findings are also consistent with studies indicating that student-centered learning approaches support students' academic engagement and learning motivation (Darling-Hammond et al., 2020).

The study also demonstrates that alternative educational approaches contribute not only to academic achievement but also to the development of problem-solving, communication, and thinking skills. In this respect, the findings provide important insights into the transition processes of students with experience in alternative education when they move to different school environments. The results suggest that different pedagogical approaches may coexist within education systems or operate across different educational stages. Such diversity in educational practices may contribute to students' multidimensional development.

4.1 Limitations of the Study

This study, designed using a qualitative approach, was conducted with a limited number of participants consisting of students, parents, and teachers from a single school implementing the BBOM model, whose graduates continued their education at the same middle school. Therefore, the findings reflect experiences within a specific educational context and cannot be generalized to all alternative education environments or to all cases of transition from alternative schools to mainstream public schools.

The findings of the study are based on qualitative interview data. Future research employing quantitative or mixed-method approaches may contribute to obtaining more comprehensive findings regarding the transition processes and developmental outcomes of students who have received education in alternative learning environments.

Implications for Educational Practice

Based on the findings of the study, several implications can be proposed.

The results of the study indicate that students who received their primary education within an alternative education model were able to adapt both academically and socially to the middle school environment. This finding suggests that student-centered learning approaches may support not only students' academic achievement but also the development of problem-solving, communication, and thinking skills. In this respect, it may be beneficial to include alternative instructional approaches more extensively in teacher education programs, particularly student-centered methods such as project-based learning, inquiry-based learning, and cooperative learning.

The findings of the study also suggest that process-oriented assessment approaches used in alternative educational environments may have positive effects on students' learning motivation and responsibility-taking skills. Therefore, it may be useful to incorporate alternative assessment methods, such as portfolio assessment, performance tasks, and project-based assignments, into mainstream school curricula in order to support process-oriented evaluation.

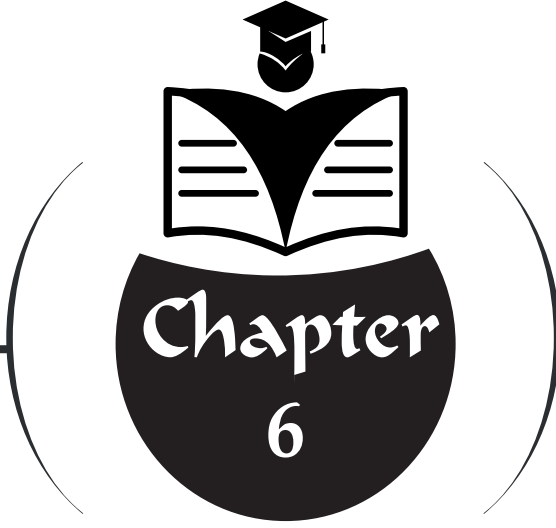
Furthermore, the results indicate that alternative educational environments may positively contribute to the development of students' communication, collaboration, and problem-solving skills. For this reason, it may be recommended that school environments provide more opportunities for students to participate in decision-making processes and engage in collaborative learning experiences.

This study focuses on a single school implementing the BBOM model. Future research may include other schools adopting the same model in order to obtain a broader understanding of the effects of alternative educational approaches. Studies conducted across different types of schools may provide more comprehensive evidence regarding the impact of alternative education practices.

Finally, future studies examining the long-term academic and social development of students who have received education in alternative learning environments may provide valuable contributions to literature. In addition, further research could include students who complete the entire K–12 educational process within alternative education models in order to better understand the long-term outcomes of such educational experiences.

References

- Biesta, G. J. J. (2015). *Good education in an age of measurement: Ethics, politics, democracy*. Boulder, CO: Paradigm Publishers.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97–140.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & MacIver, D. (1993). The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48(2), 90-101.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3), 235–266.
- Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365-379.
- Kaya, D., & Gündüz, M. (2015). Alternatif Eğitim ve Toplumsal Değişim Üzerindeki Etkisi: Waldorf Okulları Örneği. *Milli Eğitim Dergisi*, 44(205), 5-25
- Kraftl, P. (2013). *Geographies of alternative education: Diverse learning spaces for children and young people*. Bristol, UK: Policy Press.
- Lillard, A. S. (2005). *Montessori: The science behind the genius*. Oxford University Press.
- Miller, R. (2007). *What are schools for? Holistic education in American culture*. Holistic Education Press.
- Rathunde, K., & Csikszentmihalyi, M. (2005). The social context of middle school: Teachers, friends, and activities in Montessori and traditional school environments. *The Elementary School Journal*, 106(1), 59-79.
- Raywid, M. A. (1994). Alternative schools: The state of the art. *Educational Leadership*, 52(1), 26–31.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67.
- Sliwka, A. (2008). The contribution of alternative education. *Innovating to learn, learning to innovate*, 93.
- Tutal, Ö., Kaçire, İ., Atabey, E., & Milli Eğitim Bakanlığı. (2016). Öğretmen eğitiminde kullanılan alternatif öğretim yöntemlerinin akademik başarıya etkisi: Bir metaanaliz çalışması. 4. Uluslararası Eğitim Programları ve Öğretim Kongresi, Antalya, Türkiye.
- Yıldırım, A., & Şimşek, H. (2018). *Sosyal bilimlerde nitel araştırma yöntemleri* (11. baskı). Seçkin Yayıncılık, Ankara.



**THE IMPACT OF DIGITAL GAME-
BASED LEARNING IN SOCIAL
STUDIES ON THE SOCIAL-
EMOTIONAL DEVELOPMENT
OF GIFTED STUDENTS: AN AGE-
BASED ANALYSIS**

“
”

Fatih PALA¹

¹ Doç. Dr., Oltu Science and Art Center, Erzurum, Turkey, e-mail: e-r-z-u-r-u-m-25@hotmail.com, Orcid: 0000-0003-1828-0461

1. Introduction

In recent years, digital game-based learning has been increasingly recognized as an effective instructional approach that enhances students' engagement, motivation, and meaningful learning experiences [1-2]. More recent studies emphasize that digital games can also support social-emotional development, particularly when they are designed to promote collaboration, empathy, and self-regulation [3-4]. In the context of social studies education, digital games provide interactive environments that allow students to experience historical events, civic decision-making processes, and social interactions in immersive and participatory ways [5-6].

Although gifted students are often characterized by advanced cognitive abilities and high academic performance, research has shown that their social-emotional development may not always align with their intellectual growth [7-8]. Recent literature highlights that gifted learners may face challenges related to peer relationships, emotional regulation, and social adjustment, particularly during early adolescence [9-10]. Therefore, instructional approaches that intentionally address both cognitive and social-emotional dimensions are essential for supporting the holistic development of gifted students.

Social studies courses offer a particularly suitable context for fostering social-emotional learning, as they emphasize values such as citizenship, social responsibility, empathy, and ethical reasoning [11-12]. When digital game-based learning is integrated into social studies instruction, students are encouraged to adopt multiple perspectives, engage in cooperative problem-solving, and reflect on the social consequences of their decisions (13-14). Such learning experiences have been found to positively influence students' cooperation skills, emotional awareness, and social engagement [15-16].

Despite the growing body of research on digital game-based learning and gifted education, age-related differences in gifted students' social-emotional responses to digital games remain underexplored, particularly within social studies contexts [1-3]. Understanding how students at different developmental stages respond to digital game-based learning is crucial for designing developmentally responsive and inclusive instructional practices. Accordingly, this study aims to examine the impact of digital game-based learning implemented in social studies courses on the social-emotional development of gifted students, with a specific focus on age-related variations.

2. Theoretical Framework and Literature Review

Digital Game-Based Learning in Education

Digital game-based learning (DGBL) is grounded in constructivist and experiential learning theories, emphasizing active participation, problem-solving, and meaningful interaction with content [17-1]. Within this

framework, learning emerges through engagement with authentic tasks, immediate feedback, and iterative decision-making processes. These features position learners as active agents who construct knowledge through experience rather than passive recipients of information.

Beyond cognitive outcomes, recent research highlights that well-designed digital games also support motivational and affective dimensions of learning, including engagement, persistence, emotional involvement, and self-regulation [1–12]. In this sense, DGBL provides not only an instructional strategy but also a socio-emotional learning environment in which learners practice decision-making, manage emotions, and interact with others in structured yet flexible contexts.

In social studies education, DGBL has gained increasing recognition due to its capacity to simulate historical events, civic processes, and social dilemmas [5–9]. Such simulations allow learners to explore multiple perspectives, negotiate social roles, and experience the consequences of individual and collective decisions. Empirical studies indicate that digital games in social studies enhance students' conceptual understanding, critical thinking, and collaborative problem-solving skills [14], while simultaneously creating conditions conducive to social-emotional learning.

Social-Emotional Learning Framework and Digital Games

Social-emotional learning (SEL) is commonly conceptualized within the CASEL framework, which defines five core competencies: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making [12]. Contemporary educational research emphasizes that these competencies are not peripheral to academic learning but are integral to students' cognitive engagement and long-term success [15].

Building on this framework, the present study adopts a theory of change that explicitly links specific features of DGBL to distinct SEL competencies. For example, game-based scenarios requiring reflection on one's choices and emotional responses are expected to foster self-awareness, while mechanics involving goal-setting, delayed rewards, and adaptive challenges support self-management. Cooperative gameplay, role-taking, and narrative-driven perspective shifts are hypothesized to enhance social awareness and relationship skills, whereas decision-making under simulated social dilemmas aligns closely with responsible decision-making.

Recent empirical evidence supports these assumptions, demonstrating that digital games intentionally designed with social interaction, feedback, and reflection components can effectively promote SEL competencies [2]. Game environments that encourage cooperation, empathy, perspective-taking, and emotional regulation have been shown to support positive peer interaction

and emotional development [9]. Anchoring SEL outcomes within the CASEL framework therefore strengthens the conceptual coherence between theory, instructional design, measurement, and findings.

Gifted Students and Social-Emotional Development

Gifted students are typically characterized by advanced cognitive abilities, rapid learning capacity, and heightened sensitivity; however, research consistently indicates that their social-emotional development does not always progress synchronously with their intellectual growth [7–8]. This developmental asynchrony may manifest in challenges related to emotional regulation, peer relationships, and social adjustment, particularly during transitional developmental periods such as early adolescence [16].

Recent literature underscores the importance of instructional approaches that address both the cognitive and social-emotional needs of gifted learners [11]. Learning environments that promote collaboration, emotional expression, and social interaction are especially beneficial for supporting their holistic development. When thoughtfully integrated, digital games offer flexible, engaging, and socially rich contexts that align with the affective intensity and advanced cognitive engagement typical of gifted students [3].

Age as a Developmental Moderator in Digital Game-Based Learning

Developmental theory suggests that students' cognitive, emotional, and social capacities evolve significantly across age groups, influencing how they perceive, process, and respond to learning environments [1]. In the context of DGBL, age-related differences have been shown to shape learners' motivation, emotional engagement, self-regulation strategies, and patterns of social interaction [5].

From a developmental perspective, younger gifted students may benefit more from DGBL features that support emotional awareness and basic social interaction, whereas older students are more likely to engage in complex perspective-taking, moral reasoning, and collaborative decision-making. Accordingly, age is conceptualized in this study as a moderating variable that influences the strength and nature of the relationship between DGBL features and specific SEL competencies within the CASEL framework.

Despite growing interest in DGBL and gifted education, relatively few studies have systematically examined age-related differences in the social-emotional outcomes of digital game-based learning for gifted students, particularly within social studies education [7]. This gap underscores the need for developmentally responsive instructional designs that explicitly consider age as a key factor shaping SEL processes.

Research Gap and Contribution

Although prior research has documented the cognitive and motivational benefits of digital game-based learning and highlighted the importance of social-emotional development for gifted students, empirical studies that integrate DGBL, social studies education, giftedness, CASEL-based SEL competencies, and age-related differences remain limited. Addressing this gap, the present study investigates the impact of digital game-based learning implemented in social studies courses on the social-emotional development of gifted students, with a specific focus on age-related variations.

By explicitly anchoring SEL outcomes in the CASEL framework and articulating a theory-driven linkage between DGBL features and SEL competencies, the study aims to strengthen the alignment between theory, measurement, and findings. In doing so, it contributes evidence-based insights into developmentally responsive, inclusive, and technology-supported instructional practices in social studies education.

3. Method

This study employs a mixed-methods research design to investigate the effects of digital game-based learning in social studies courses on the social-emotional development of gifted students across different age groups. Mixed-methods research enables the integration of quantitative and qualitative data, providing a more comprehensive understanding of complex educational phenomena [18].

The quantitative component is based on a quasi-experimental pretest-posttest control group design. Gifted students are assigned to either an experimental group receiving digital game-based learning instruction or a control group receiving traditional instruction. This design allows for the examination of instructional effects while accounting for pre-existing differences between groups [19].

Age is treated as a comparative variable, and participants are grouped according to developmental stages to examine age-related differences in social-emotional outcomes. The qualitative component follows a case study approach, incorporating classroom observations, student reflections, and semi-structured interviews to deepen the interpretation of quantitative findings [20]. The integration of both data types ensures methodological triangulation and enhances the validity of the findings [21].

Participants

The participants of this study consist of gifted students enrolled in formal and supplementary educational programs for gifted education. Participants were selected using purposive sampling, as the study specifically focuses

on students identified as gifted through official assessment and placement procedures [22].

To examine age-related differences, students were grouped according to their developmental stages. The sample included students from three different age groups (e.g., 9–10, 11–12, and 13–14 years). Within each age group, participants were assigned to either an experimental group, which engaged in digital game-based learning activities in social studies courses, or a control group, which received instruction through traditional teaching methods.

Participation in the study was voluntary, and informed consent was obtained from both students and their parents. Ethical approval was secured prior to data collection, and all procedures were conducted in accordance with established ethical guidelines for educational research [23].

Data Collection Instruments

Multiple data collection instruments were employed to examine the effects of digital game-based learning on the social-emotional development of gifted students. Using both quantitative and qualitative tools allowed for data triangulation and enhanced the validity of the findings [18].

Quantitative Data Collection Instrument

The quantitative data of the study were collected using a Social-Emotional Learning (SEL) Scale developed by the researcher. The scale was designed to assess the social-emotional development levels of gifted students in digital game-based learning environments. In the scale development process, the theoretical framework of social-emotional learning and current studies in the literature were taken as a basis, with particular emphasis on the dimensions of self-awareness, self-regulation, social awareness, relationship skills, and responsible decision-making [15-16].

During the scale development process, a comprehensive review of the relevant literature was conducted, and an initial item pool was created in accordance with the identified dimensions. The items were constructed by considering the cognitive, emotional, and developmental characteristics of gifted students, as well as their age levels. To ensure content and face validity, the draft scale was reviewed by academic experts in the fields of social studies education, educational measurement and evaluation, and gifted education. Revisions were made based on expert feedback.

The developed Social-Emotional Learning Scale was administered as a pretest and posttest to determine the effects of digital game-based learning practices on students' social-emotional development. The scale was applied to both experimental and control groups at the beginning of the research process and after the completion of the instructional intervention, allowing

for a quantitative comparison of changes between groups and across age levels.

The collected data were analyzed in accordance with the psychometric properties of the scale, and the effects of digital game-based learning on the social-emotional development of gifted students were examined within the context of age as a comparative variable.

Qualitative Data Collection Instruments

The qualitative data of the study were collected through classroom observations, student reflective forms, and semi-structured interviews. These qualitative data collection instruments were implemented within the context of digital game-based learning activities designed in direct alignment with the learning objectives of the social studies curriculum. The activities were structured to address the core aims of social studies education, including understanding and interpreting historical events, developing citizenship awareness, proposing solutions to social issues, and engaging in democratic decision-making processes.

Classroom observations focused on students' interaction patterns, emotional responses, and collaborative behaviors during the digital game-based learning process. Through these observations, indicators of students' social-emotional development—such as empathy, responsibility, group-based decision-making, and social awareness—were systematically examined within a natural learning environment. Particular attention was given to the roles students assumed within game scenarios and their responses to historical and social situations presented in the games.

Student reflective forms enabled learners to express their perceptions, emotional experiences, and social interactions related to social studies-based digital game activities in their own words. These reflections were used to examine students' evaluations of the contributions of game-based activities to their social-emotional development and their attitudes toward the social studies course.

Semi-structured interviews were conducted to obtain more in-depth and detailed insights into students' social-emotional development processes. The interviews explored the effects of digital game-based social studies activities on students' empathy, collaboration, and responsible decision-making skills, as well as how age-related differences in learning and emotional experiences emerged throughout the process [20].

The combined use of quantitative and qualitative data collection instruments allowed for a multidimensional examination of the effects of digital game-based learning on the social-emotional development of gifted students within the context of social studies education, supporting a holistic interpretation of the findings through multiple data sources.

Data Analysis

The data obtained in this study were analyzed using both quantitative and qualitative analysis techniques in accordance with the mixed-methods research design. This integrative approach allowed for a comprehensive examination of the effects of digital game-based learning on the social-emotional development of gifted students across different age groups [18].

Quantitative Data Analysis

Quantitative data collected through the researcher-developed Social-Emotional Learning (SEL) Scale were analyzed using statistical software. Prior to analysis, data were screened for missing values and normality assumptions. Descriptive statistics were calculated to determine participants' overall social-emotional development levels. To examine the effects of digital game-based learning, pretest and posttest scores of the experimental and control groups were compared. Depending on the distribution of the data, paired-samples t-tests, independent-samples t-tests, and analysis of variance (ANOVA) were conducted. Age was treated as a comparative variable to identify age-related differences in social-emotional development. Effect sizes were calculated to determine the magnitude of the observed effects.

Qualitative Data Analysis

Qualitative data obtained from classroom observations, student reflective forms, and semi-structured interviews were analyzed using thematic analysis. The analysis followed a systematic process involving data familiarization, initial coding, theme development, and interpretation [24]. Codes were generated based on recurring patterns related to social-emotional development, such as empathy, collaboration, responsibility, emotional regulation, and social awareness. The identified themes were examined in relation to social studies content and age-related differences.

Integration of Quantitative and Qualitative Data

Findings from the quantitative and qualitative analyses were integrated during the interpretation phase to ensure methodological triangulation. Qualitative findings were used to explain and enrich quantitative results, thereby providing a more holistic understanding of how digital game-based learning influences the social-emotional development of gifted students within the context of social studies education [21].

Research Process

The research process was carried out in a systematic and sequential manner in line with the mixed-methods research design. The study consisted of preparation, implementation, data collection, and analysis stages.

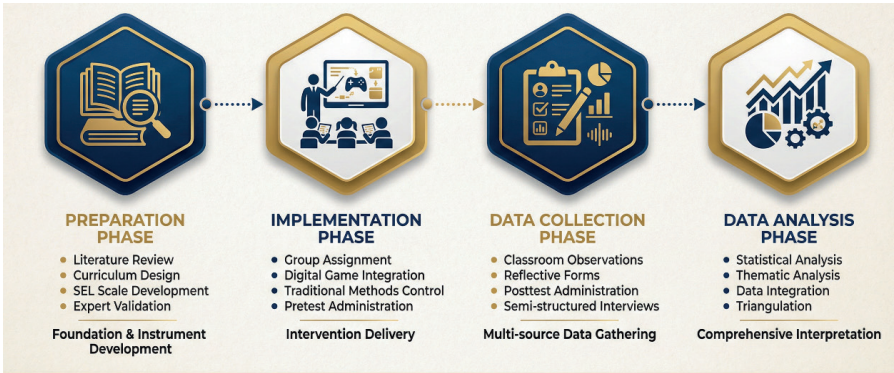
In the preparation phase, a comprehensive review of the literature on digital game-based learning, social-emotional learning, gifted education, and social studies education was conducted. Based on the theoretical framework and research objectives, digital game-based learning activities aligned with the social studies curriculum were designed. Simultaneously, the quantitative data collection instrument—the researcher-developed Social-Emotional Learning (SEL) Scale—was developed, and qualitative data collection tools (observation forms, reflective forms, and interview protocols) were prepared. Expert feedback was obtained to ensure content validity and clarity of the instruments.

During the implementation phase, participants were assigned to experimental and control groups. The experimental group participated in digital game-based learning activities integrated into social studies lessons, while the control group received instruction through traditional teaching methods. The intervention was conducted over a period of several weeks within regular class hours. Prior to the instructional intervention, the SEL Scale was administered as a pretest to both groups.

In the data collection phase, qualitative data were gathered throughout the intervention through classroom observations and student reflective forms, allowing for continuous monitoring of students' social-emotional behaviors and experiences. Following the completion of the instructional process, the SEL Scale was administered as a posttest to both experimental and control groups. Semi-structured interviews were then conducted with selected participants to obtain in-depth insights into students' social-emotional development and age-related learning experiences.

In the data analysis phase, quantitative and qualitative data were analyzed separately using appropriate statistical and thematic analysis techniques. The results were subsequently integrated during the interpretation stage to ensure triangulation and to provide a comprehensive understanding of the effects of digital game-based learning on the social-emotional development of gifted students within the context of social studies education.

Figure 1. *Research Process*



Validity and Reliability

To ensure the scientific rigor of the study, multiple strategies were employed to establish the validity and reliability of both quantitative and qualitative data.

The validity of the researcher-developed Social-Emotional Learning (SEL) Scale was established through content and construct validity procedures. Content validity was ensured by consulting experts in social studies education, educational measurement and evaluation, and gifted education. Based on expert feedback, revisions were made to improve the clarity, relevance, and representativeness of the scale items.

Construct validity was examined using exploratory factor analysis (EFA) to identify the underlying factor structure of the scale. The factor structure was found to be consistent with the theoretical dimensions of social-emotional learning, including self-awareness, self-regulation, social awareness, relationship skills, and responsible decision-making.

The reliability of the scale was assessed using Cronbach’s alpha coefficient. The overall reliability coefficient and the reliability values for each sub-dimension indicated acceptable to high internal consistency, demonstrating that the scale reliably measures students’ social-emotional development.

The trustworthiness of qualitative data was ensured through strategies addressing credibility, transferability, dependability, and confirmability. Credibility was enhanced through prolonged engagement in the research setting and the use of multiple qualitative data sources, including observations, reflective forms, and interviews. Transferability was supported by providing detailed descriptions of the research context, participants, and procedures.

Dependability and confirmability were strengthened through systematic data collection procedures and transparent documentation of the analysis process. Coding decisions and emerging themes were reviewed to ensure

consistency, and qualitative findings were supported by direct evidence from multiple data sources.

Finally, methodological triangulation was achieved by integrating quantitative and qualitative findings during the interpretation phase. The convergence of results from different data sources enhanced the overall validity of the study and provided a more comprehensive understanding of the effects of digital game-based learning on the social-emotional development of gifted students in social studies education.

4. Findings

Table 1. Comparison of Pretest and Posttest Social-Emotional Learning Scores

| Group | Test | N | Mean (M) | Std. Deviation (SD) | t | p |
|--------------|----------|----|----------|---------------------|------|-------|
| Experimental | Pretest | 30 | 3.21 | 0.42 | | |
| Experimental | Posttest | 30 | 4.05 | 0.38 | 6.84 | .001* |
| Control | Pretest | 30 | 3.18 | 0.40 | | |
| Control | Posttest | 30 | 3.25 | 0.41 | 1.12 | .268 |

* $p < .05$

As presented in Table 1, the pretest results indicated that there was no statistically significant difference between the experimental and control groups in terms of social-emotional learning scores prior to the intervention. This finding suggests that both groups were comparable at the beginning of the study.

Following the implementation of digital game-based learning activities, the experimental group demonstrated a statistically significant increase in posttest social-emotional learning scores. The observed difference between pretest and posttest scores for the experimental group was found to be significant ($t = 6.84, p < .05$), indicating a substantial improvement in students' social-emotional development.

In contrast, the control group showed no statistically significant change between pretest and posttest scores ($p > .05$). This finding suggests that traditional instructional practices did not result in a meaningful improvement in social-emotional learning outcomes over the same period.

Table 2. Pretest–Posttest Comparisons of SEL Sub-Dimensions (Experimental Group)

| SEL Sub-Dimension | Test | N | Mean (M) | Std. Deviation (SD) | t | p |
|-------------------|----------|----|----------|---------------------|------|-------|
| Self-Awareness | Pretest | 30 | 3.28 | 0.44 | | |
| | Posttest | 30 | 4.12 | 0.36 | 6.21 | .001* |
| Self-Regulation | Pretest | 30 | 3.10 | 0.47 | | |
| | Posttest | 30 | 3.95 | 0.40 | 5.87 | .001* |

| SEL Sub-Dimension | Test | N | Mean (M) | Std. Deviation (SD) | t | p |
|-----------------------------|----------|----|----------|---------------------|------|-------|
| Social Awareness | Pretest | 30 | 3.18 | 0.43 | 7.04 | .001* |
| | Posttest | 30 | 4.20 | 0.35 | | |
| Relationship Skills | Pretest | 30 | 3.25 | 0.41 | 6.76 | .001* |
| | Posttest | 30 | 4.18 | 0.37 | | |
| Responsible Decision-Making | Pretest | 30 | 3.23 | 0.45 | 5.42 | .001* |
| | Posttest | 30 | 4.00 | 0.39 | | |

*p < .05

The results presented in Table 2 indicate that digital game-based learning activities integrated into social studies instruction led to statistically significant improvements across all social-emotional learning sub-dimensions for the experimental group.

Among the sub-dimensions, the largest increase was observed in social awareness, suggesting that game scenarios grounded in historical and social contexts effectively supported students’ ability to understand others’ perspectives, develop empathy, and recognize social dynamics. This finding is particularly consistent with the aims of social studies education, which emphasize citizenship, empathy, and social responsibility.

Significant gains were also found in relationship skills and self-awareness, indicating that collaborative gameplay and role-based scenarios encouraged students to reflect on their own emotions while engaging constructively with peers. Improvements in self-regulation suggest that game-based tasks requiring strategic thinking and delayed gratification supported students’ emotional control and persistence.

Although responsible decision-making showed comparatively smaller gains, the difference was still statistically significant, implying that repeated exposure to decision-based game mechanics contributed positively to students’ ability to evaluate consequences within social and civic contexts.

Table 3. *Posttest SEL Scores of Experimental and Control Groups*

| Group | N | Mean (M) | Std. Deviation (SD) | t | p |
|--------------------|----|----------|---------------------|------|-------|
| Experimental Group | 30 | 4.09 | 0.38 | 7.24 | .001* |
| Control Group | 30 | 3.31 | 0.42 | | |

*p < .05

The independent samples t-test results presented in Table 3 reveal a statistically significant difference between the posttest social-emotional learning scores of students in the experimental and control groups. Students who participated in digital game-based learning activities integrated into social

studies instruction demonstrated significantly higher levels of social-emotional development compared to those who received traditional instruction.

This finding indicates that the observed improvement in the experimental group cannot be attributed solely to maturation or routine classroom experiences. Rather, it suggests that digital game-based learning served as an effective instructional intervention in fostering social-emotional competencies among gifted students.

The higher posttest scores of the experimental group highlight the role of interactive, scenario-based, and collaborative digital games in supporting emotional engagement, peer interaction, and reflective decision-making—core components of both social-emotional learning and social studies education. In contrast, the control group, which followed conventional instructional approaches, showed more limited gains in these areas.

Table 4. *Posttest Social-Emotional Learning (SEL) Scores by Age Groups in the Experimental and Control Groups*

| Group | Age Group | N | Mean (M) | Standard Deviation (SD) |
|---------------------------|-------------|-----------|-------------|-------------------------|
| Experimental | 9–10 years | 10 | 3.78 | 0.41 |
| Experimental | 11–12 years | 10 | 4.12 | 0.36 |
| Experimental | 13–14 years | 10 | 4.38 | 0.34 |
| Experimental Total | — | 30 | 4.09 | 0.38 |
| Control | 9–10 years | 10 | 3.54 | 0.45 |
| Control | 11–12 years | 10 | 3.60 | 0.43 |
| Control | 13–14 years | 10 | 3.72 | 0.40 |
| Control Total | — | 30 | 3.62 | 0.44 |

Note. The experimental and control groups had equal sample sizes ($N = 30$). Age-based analyses were conducted with $n = 10$ students in each age group. Values reported in the table represent posttest SEL scores.

As shown in Table 4, across all age groups, students in the experimental group obtained higher posttest social-emotional learning (SEL) mean scores than their counterparts in the control group. Within the experimental group, SEL scores demonstrated a consistent increase with age, with the lowest mean observed in the 9–10 age group ($M = 3.78$, $SD = 0.41$) and the highest mean in the 13–14 age group ($M = 4.38$, $SD = 0.34$).

In the control group, a more limited age-related increase was observed. Mean SEL scores increased from $M = 3.54$ ($SD = 0.45$) in the 9–10 age group to $M = 3.72$ ($SD = 0.40$) in the 13–14 age group. However, this increase was notably smaller compared to the pattern observed in the experimental group.

Additionally, the decreasing standard deviation values across age groups in the experimental condition indicate that students' scores became more

homogeneous as age increased. This pattern suggests that the instructional approach implemented in the experimental group produced more consistent social-emotional learning outcomes, particularly among older students.

Table 5. ANOVA Results for Posttest SEL Scores by Age Groups (Experimental Group)

| Source of Variation | SS | df | MS | F | p |
|---------------------|----|----|----|------|-------|
| Between Groups | – | 2 | – | 8.96 | .001* |
| Within Groups | – | 57 | – | | |
| Total | – | 59 | | | |

*p < .05

The ANOVA results presented in Table 5 indicate that there was a statistically significant difference in posttest SEL scores among the three age groups, $F(2, 57) = 8.96, p < .05$. This finding suggests that age had a significant effect on social-emotional learning outcomes in the experimental group. In other words, the observed increase in mean SEL scores from the 9–10-year-old group ($M = 3.78$) to the 11–12-year-old group ($M = 4.12$) and the 13–14-year-old group ($M = 4.38$) is unlikely to have occurred by chance. These results support the conclusion that older participants in the study demonstrated higher levels of SEL, highlighting the potential influence of developmental factors on social-emotional learning.

Development of Empathy and Perspective-Taking Through Digital Game-Based Learning

Theme 1: Empathy and Perspective-Taking

As a result of the qualitative data analysis, it was found that digital game-based Social Studies activities significantly supported gifted students’ abilities to develop empathy and understand multiple perspectives. In particular, games incorporating historical and social scenarios encouraged students not only to make knowledge-based decisions but also to evaluate the humanistic and societal consequences of those decisions.

During classroom observations, students were seen to assume different social roles throughout the gameplay process and to use expressions aimed at understanding others’ emotions and thoughts through these roles. It was observed that students demonstrated a tendency to “put themselves in others’ shoes” when confronted with moral dilemmas in game scenarios and considered alternative perspectives during group discussions.

Analysis of student reflective evaluation forms revealed that students in the middle and upper age groups expressed empathy-related statements in a deeper and more well-reasoned manner. These students emphasized that

digital games provided them with opportunities to understand the emotions experienced by historical figures or social groups.

Data obtained from semi-structured interviews indicated that digital game-based learning supported empathy and perspective-taking skills at varying levels depending on age. While younger students tended to express empathy primarily through emotional reactions, older students approached empathy in terms of responsible decision-making and anticipating social consequences.

Theme 2: Empathy and Perspective-Taking

The results of the qualitative analysis indicate that digital game-based Social Studies activities significantly enhanced gifted students' empathy and perspective-taking skills. In particular, game scenarios designed around historical events and social issues enabled students to understand the emotions, thoughts, and living conditions of different individuals and groups.

Classroom observations revealed that students demonstrated a tendency to "think from another person's position" in line with the roles they assumed during the gameplay process. During group discussions, students evaluated the social consequences of their decisions and took alternative perspectives into account.

Student statements related to this process were particularly noteworthy:

"I played the role of a villager in the game. I thought about what the decision to go to war meant not only for the rulers, but also for the people." (S13, age 13)

"Normally in history class we just learn about events, but in the game I tried to understand how people felt." (S7, age 11)

Semi-structured interviews further demonstrated that empathy skills deepened with age. While younger students tended to express empathy primarily at an emotional level, older students approached empathy in terms of responsible decision-making and anticipating social consequences.

Improvement in Collaboration and Social Interaction Skills

Theme 3: Collaboration and Social Interaction

Qualitative data revealed that digital game-based learning activities strengthened students' collaboration and social interaction skills. Throughout the gameplay process, students communicated toward shared goals, engaged in task distribution, and actively participated in group decision-making processes.

Classroom observations indicated that, particularly in multiplayer and problem-solving-based games, students exchanged ideas with one another and attempted to resolve conflicts through negotiation.

Statements from student reflective forms further supported this finding:

“I had to think as a group, not on my own. My friends’ ideas changed my decision.” (S19, age 12)

“If we didn’t listen to everyone’s ideas, we couldn’t win the game.” (S4, age 10)

This finding demonstrates that digital game-based Social Studies instruction supports not only individual learning but also social interaction and democratic participation skills.

Development of Responsible Decision-Making and Self-Regulation

Theme 4: Responsible Decision-Making and Self-Regulation

Another important finding of the study is that the digital game-based learning process supported students’ responsible decision-making and self-regulation skills. Within the game scenarios, students encountered situations involving limited resources, time pressure, and ethical dilemmas, which required them to make thoughtful decisions and anticipate potential outcomes.

Students described this process as follows:

“If I made a wrong decision, everyone in the game was affected, so I didn’t rush.” (S22, age 14)

“I learned to stay calm in the game, because when I made decisions too quickly, I made mistakes.” (S9, age 11)

These statements indicate that digital game-based learning contributed to the development of students’ ability to manage emotional responses and engage in conscious, reflective decision-making.

Integration of Quantitative and Qualitative Findings

When the quantitative and qualitative findings are considered together, it is evident that digital game-based Social Studies instruction significantly supports the social-emotional development of gifted students. Age-related differences identified in the quantitative analyses are explained by the qualitative findings, which indicate that empathy, collaboration, and responsible decision-making skills deepen as students grow older.

These integrated findings demonstrate that digital game-based learning offers a developmentally responsive, interaction-oriented, and inclusive instructional approach in Social Studies education.

5. Discussion

The present study investigated the effects of digital game-based learning (DGBL) in social studies on the social-emotional development of gifted

students, with a particular focus on age-based differences. The findings indicated significant improvements in social-emotional learning (SEL) outcomes for students engaged in digital game-based activities, with older age groups showing greater gains. In addition, thematic qualitative findings underscored increases in empathy, collaboration, and responsible decision-making. These results align with and extend current research evidence on the educational potential of game-based learning interventions.

Consistent with meta-analytic evidence showing that game-based learning can positively impact social and emotional development, our results demonstrate that structured DGBL activities can enhance students' social competencies [25]. Systematic reviews indicate that game-based learning has a moderate to large effect on social and emotional development across age groups [26], and the current study's quantitative findings support these conclusions in the context of social studies education by showing statistically significant improvements in overall SEL scores and sub-dimensions such as social awareness and relationship skills.

The age-based differences observed in this study are also meaningful in light of existing literature suggesting that developmental factors influence how learners engage with and benefit from interactive learning environments [27]. Older students' greater gains may reflect increased cognitive complexity and reflective capacity that enable deeper engagement with scenario-based tasks, particularly those involving historical and civic challenges within social studies content. This observation is consistent with research in related educational contexts showing that age and developmental readiness moderate the effects of instructional technology on learning outcomes [28].

Qualitative themes identified in this study—especially empathy, collaboration, and responsible decision-making—mirror findings from recent research examining digital interventions designed to support SEL skills. For instance, mixed-methods studies using digital game environments report enhanced empathy and intercultural competence among participants exposed to interactive scenarios [29]. These parallels suggest that digital games can function as rich, immersive contexts for promoting not only cognitive learning but also social and emotional competence when appropriately aligned with curricular goals and developmentally responsive design.

Furthermore, the integration of SEL into subject-area instruction, particularly social studies, reflects broader educational priorities emphasizing 21st-century skills, such as empathy, critical thinking, and collaboration. Studies on digitally mediated SEL interventions during and after the pandemic have highlighted the importance of technology in addressing emotional and social learning needs, especially when traditional classroom interaction is limited [30]. Our findings extend this work by demonstrating that DGBL can

be similarly effective within subject-specific curricular frameworks, reinforcing the idea that SEL and academic content can be taught synergistically rather than in isolation.

However, the current study also complements literature on potential complexities associated with technology use. Although not a focus of this research, some studies suggest that digital game usage may have negative correlates—such as excessive screen time or addictive behaviors that could moderate positive outcomes (e.g., digital game addiction negatively correlating with SEL in other contexts) [5]. These considerations underscore the importance of purposeful instructional design and moderation when applying digital games in educational settings.

In summary, the study's results converge with recent research acknowledging the pedagogical efficacy of game-based learning for promoting social and emotional development, while also highlighting age-related differences and content-specific applications in social studies. These findings contribute to a growing body of evidence that DGBL can support meaningful learning outcomes beyond academic achievement when thoughtfully integrated with SEL frameworks and disciplinary objectives.

6. Conclusion and Implications

This study examined the effects of digital game-based learning integrated into social studies instruction on the social-emotional development of gifted students, with a particular emphasis on age-related differences. The findings revealed that digital game-based learning had a significant and positive impact on students' overall social-emotional learning outcomes as well as on specific sub-dimensions, including empathy, collaboration, self-regulation, and responsible decision-making.

Quantitative results demonstrated that students in the experimental group outperformed those in the control group, indicating that the observed improvements were attributable to the instructional intervention rather than to routine classroom experiences. Moreover, age-based analyses showed that while digital game-based learning supported social-emotional development across all age groups, its effects were more pronounced among older gifted students. These findings suggest that developmental readiness plays an important role in maximizing the benefits of game-based learning environments.

Qualitative findings further enriched the quantitative results by revealing how students experienced these gains in practice. Through immersive scenarios, role-based decision making, and collaborative problem-solving tasks, students demonstrated increased empathy, deeper perspective-taking, and improved social interaction skills. Together, the integrated findings provide strong evidence that digital game-based learning constitutes an

effective, developmentally responsive, and inclusive instructional approach for supporting the social-emotional development of gifted learners in social studies education.

Implications for Practice

The findings of this study offer several important implications for educational practice. First, social studies teachers are encouraged to integrate purposefully designed digital games into their instruction, particularly games that incorporate historical dilemmas, civic decision-making, and collaborative problem-solving. Such designs not only support curricular learning outcomes but also foster essential social-emotional competencies.

Second, instructional designers and educators should consider age-appropriate complexity when implementing digital game-based learning. Younger students may benefit from emotionally accessible and guided scenarios, whereas older gifted students may engage more deeply with complex narratives involving ethical dilemmas and societal consequences.

Third, the results highlight the importance of viewing social-emotional learning not as an isolated component but as an integrated element of subject-area teaching. Embedding SEL objectives within social studies content can enhance students' engagement while promoting empathy, citizenship awareness, and responsible participation in social life.

Implications for Research

From a research perspective, this study contributes to the growing literature on digital game-based learning by demonstrating its relevance for social-emotional development and gifted education within a subject-specific context. Future studies may expand on this work by employing longitudinal designs to examine the long-term effects of digital game-based learning on social-emotional competencies.

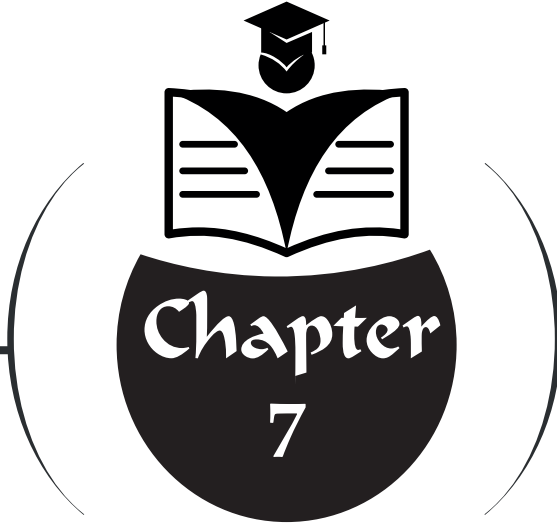
Additionally, further research could explore the role of other moderating variables—such as gender, cultural context, or prior gaming experience—to better understand how different learner characteristics influence outcomes. Comparative studies involving different subject areas may also help clarify the unique contribution of social studies to social-emotional learning through digital game-based approaches.

In conclusion, the findings of this study underscore the potential of digital game-based learning as a powerful pedagogical tool that bridges academic content and social-emotional development. When thoughtfully designed and developmentally aligned, digital games can transform social studies classrooms into meaningful learning environments that support both the cognitive and emotional growth of gifted students.

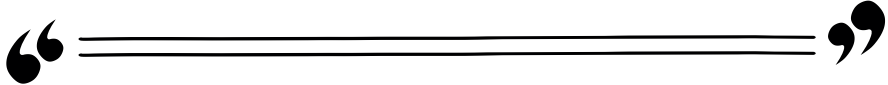
References

- [1] Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258-283.
- [2] Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in human behavior*, 63, 50-58.
- [3] Boghian, I., & Cojocariu, V. M. (2023). Using games to build social emotional learning skills. *Revista Romaneasca Pentru Educatie Multidimensionala*, 15(1), 622-656.
- [4] Alotaibi, M. S. (2024). Game-based learning in early childhood education: a systematic review and meta-analysis. *Frontiers in psychology*, 15, 1307881.
- [5] Squire, K. (2013). Video games and learning: Teaching and participatory culture in the digital age. *Alberta Journal of Educational Research*, 59(1), 129-132.
- [6] Cadiz, G. S., Lacre, G. J. R., Delamente, R. L., & Diquito, T. J. A. (2023). Game-based learning approach in science education: a meta-analysis. *International Journal of Social Science and Human Research*, 6(03), 1856-1865.
- [7] Neihart, M. (2021). *The social and emotional development of gifted children: What do we know?*. Routledge.
- [8] Cross, T. L. (2021). *On the social and emotional lives of gifted children*. Routledge.
- [9] Peterson, J. S., Assouline, S. G., & Jen, E. (2021). Responding to concerns related to the social and emotional development of gifted adolescents. In *The handbook of secondary gifted education* (pp. 65-90). Routledge.
- [10] Papadopoulos, D. (2020). Psychological framework for gifted children's cognitive and socio-emotional development: A review of the research literature and implications. *Journal for the Education of Gifted Young Scientists*, 8(1), 305-323.
- [11] Swan, K., Barton, K. C., Buckles, S., Burke, F., Charkins, J., Grant, S. G., ... & Wiesner-Hanks, M. (2013). The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K-12 civics, economics, geography, and history.
- [12] Dwomoh, R. K., & Johnson, C. S. (2025). Fostering Civic Readiness: Assessing the Scope and Content of Civics in Middle-Level Social Studies Curricula. *The Social Studies*, 1-31.
- [13] Squire, K., & Jenkins, H. (2003). Harnessing the power of games in education. *Insight*, 3(1), 5-33.
- [14] Wang, X. M., Wang, S. M., Wang, J. N., Hwang, G. J., & Xu, S. (2023). Effects of a two-tier test strategy on students' digital game-based learning performances and flow experience in environmental education. *Journal of Educational Computing Research*, 60(8), 1942-1968.
- [15] Durlak, J. A. (Ed.). (2015). *Handbook of social and emotional learning: Research and practice*. Guilford Publications.

- [16] Domitrovich, C., Li, Y., Kendziora, K., & Greenberg, M. T. (2025). Impact of schoolwide support for social and emotional learning: The effect of the CASEL school guide on organizational capacity for SEL and systemic SEL. *Social and Emotional Learning: Research, Practice, and Policy*, 100156.
- [17] Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1), 20-20.
- [18] Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.
- [19] Shadish, W. R. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage Learning.
- [20] Yin, R. K. (2018). *Case study research and applications (Vol. 6)*. Thousand Oaks, CA: Sage.
- [21] Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), 112-133.
- [22] Fraenkel, W., & Wallen, & N. E. Hyun, (2012). *How to design and evaluate research in education, (Eight Edition)*. New York: McGraw-Hill.
- [23] Banks, J. A. (2016). Expanding the epistemological terrain: Increasing equity and diversity within the American Educational Research Association. *Educational Researcher*, 45(2), 149-158.
- [24] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- [25] Alotaibi, M. S. (2024). Game-based learning in early childhood education: a systematic review and meta-analysis. *Frontiers in psychology*, 15, 1307881.
- [26] Liverman, E., Antognoli, D., Elaiho, C., McGuire, M., Stoltenburg, A., Navarrete, A., ... & Meurer Sr, J. (2025). Game-Based Social-Emotional Learning for Youth: School-Based Qualitative Analysis of Brain Agents. *JMIR Formative Research*, 9, e67550.
- [27] Bali, M. M. E. I., Najiburrahman, Fathony, A., Salma, Maghfirah, E., & Farida, L. A. (2021, May). Utilization of Zoom Cloud in M3D (Maze 3D) Game-Based Learning to Develop Early Childhood Social-Emotional Skills. In *IOP Conference Series: Materials Science and Engineering (Vol. 1125, No. 1, p. 012061)*. IOP Publishing.
- [28] Elsheikh, A., Othman, A., & Al-Thani, D. (2025). Systematic analysis of cutting-edge technology for the wellbeing and safety of older persons. *Disability and Rehabilitation: Assistive Technology*, 20(8), 2691-2724.
- [29] Mukund, V., Sharma, M., Srivastva, A., Sharma, R., Farber, M., & Chatterjee Singh, N. (2022). Effects of a digital game-based course in building adolescents' knowledge and social-emotional competencies. *Games for Health Journal*, 11(1), 18-29.
- [30] Warren, J. L. (2023). Digital and interactive technologies for children's mental health and socio-emotional wellbeing: Exploring potential, gaps, and design opportunities.



SCENARIO BASED LEGO ROBOTICS IN SCIENCE TEACHER EDUCATION: THE CASE OF WATER THEME¹



Türkan Aybike AKARCA²

Mustafa ERGUN³

1 This study was produced from the first author's master's thesis under the supervision of the second author.

2 M. Sc., Ondokuz Mayıs University, Graduate School of Education, Department of Science Education <https://orcid.org/0000-0003-2155-7011>

3 Assoc. Prof. Dr., Ondokuz Mayıs University, Faculty of Education, Department of Science Education, Samsun, Türkiye <https://orcid.org/0000-0003-4471-6601>

Science education in developed countries is continually updated to meet the demands of the era, improve education systems, and keep pace with scientific advancement. In this context, the questions of what the objectives of science education should be and how these objectives can be achieved remain important for both curriculum development and pedagogical innovation. In recent years, responses to these questions have increasingly been shaped by the STEM (Science, Technology, Engineering, Mathematics) approach and by educational robotics (ER) applications as one of its components.

Although robotic technologies have long been used in various fields, the widespread use of ER and coding based learning environments in education is relatively recent. It is suggested that ER can make significant contributions at all levels of education (Johnson, 2003). In this regard, robotics is seen as a tool that integrates science and technology, making learning more effective and lasting (Wood, 2003). Indeed, while some studies reveal that robotics supported learning environments improve students' problem solving skills (Silik, 2016), others emphasize that these environments enrich learning processes and play a significant role in education (Koç-Şenol & Büyük, 2015). Additionally, international organizations report that the number of robots used in education and entertainment is rapidly increasing (Barreto & Benitti, 2012).

Researchers who adopt a constructivist approach to learning argue that robotic applications contribute significantly to learning by increasing students' active participation (Papert, 1993). However, it has also been noted that experimental studies on the effects of robotics at the K-12 level are limited (Williams et al., 2007). Today, this field has developed rapidly, and ER has become an important component of educational research. In particular, numerous studies have been conducted from primary school to university level using LEGO Mindstorms kits, and these tools are widely utilized in education (Bilotta et al., 2009).

Although various tools exist in the field of ER, LEGO robots are preferred in the majority of studies (Benitti, 2012, as cited in Aydın, 2018). In particular, the LEGO Mindstorms EV3 kit is widely used in international competitions and educational applications and plays an important role in developing students' robotic coding skills. Similarly, robotics competitions are organized in Türkiye, and interest in this field is steadily increasing.

In light of these developments, increasing studies on the use of ER in science education and supporting them with practice based research have become important. In particular, conducting studies aimed at developing PSTs' engineering, problem solving, and creativity skills is considered critical for educating qualified individuals of the future. In this context, the question of how science concepts can be taught through scenario based approaches and ER, based on pre-service teachers' existing skills, constitutes the starting point of this study.

Robotics in Teacher Education: Skills, Attitudes, and Challenges

ER is considered an important tool in education due to its structure that integrates disciplines such as mathematics, physics, engineering, and programming (Johnson, 2003). Studies in this field indicate that robotic applications support students' learning processes and are particularly effective in problem solving and hands-on learning (Jung, 2013; Altin & Pedaste, 2013). Research conducted at different educational levels has shown that robotics applications increase students' interest in engineering and improve their teamwork and problem-solving skills (Karp et al., 2010; Cejka, Rogers, & Portsmore, 2006). Furthermore, these applications are reported to support students' technological and social development (Bilotta et al., 2009). It has also been emphasized that more comprehensive learning environments can be created by enriching LEGO kits with additional tools (Klassner & Anderson, 2003). In the context of teacher education, it is stated that teachers need support for robotics applications to be effective (Johnson, 2003). These environments have been found to enhance pre-service teachers' problem solving and collaboration skills (Silik, 2016) and to positively influence their scientific process skills and attitudes toward science teaching (Okkesim, 2014). However, it is understood that studies in the field of ER in Türkiye are limited, and given the importance of STEM and 21st-century skills, there is a need for further research in this area.

The aim of the study

This research investigates the integration of LEGO robotics into a scenario based learning environment to facilitate the acquisition of water related scientific concepts among pre-service science teachers (PST). Furthermore, it evaluates how individual competencies, including scientific reasoning, creativity, and decision making, contribute to the overall effectiveness of this technology-enhanced pedagogical approach. Additionally, the study explores the progression of PSTs' coding competencies alongside their qualitative perspectives on scenario based instruction and LEGO robotics integration. By analyzing these viewpoints, the research seeks to determine the perceived impact of robotic tools on the overall engagement and technical development of future educators.

Based on the identified research problem, the sub-problems are as follows:

1. How do PSTs' scientific process skills, scientific creativity, and decision making skills affect their coding achievement in a scenario and LEGO robotics supported science learning environment?
2. What are PSTs' views on a scenario based, educationally supportive robotics learning environment?
3. How do scenarios and a LEGO robotics supported learning environment contribute to PSTs' demonstration of scientific creativity, problem solving, collaboration, and 21st-century skills?

The methodology of the study

Case study research examines modern day events in their actual settings using multiple types of evidence, particularly when the subject and its context are deeply intertwined (Yin, 1984, as cited in Yıldırım & Şimşek, 2016). Based on this framework in the literature, a holistic single case design was employed in this study.

All twelve participating PSTs possess a LEGO coding certification, ensuring a baseline level of technical proficiency for the research. These participants were organized into pairs to implement, with each group’s involvement spanning approximately 6 hours. During this period, the first 1.5 hours were allocated to the data collection instruments required for conducting the preliminary analyses, while the final hour was devoted to the data collected at the end of the study.

Both quantitative and qualitative data were collected in this study. These included the Scientific Process Skills Test (SPST), originally developed by Burns, Okey, and Wise (1985) and adapted into Turkish by Geban, Aşkar, and Özkan (1992). The Scientific Creativity Test (SCT), developed by Hu and Adey (2002) and the Decision Making Skills Test (DMST), developed by Bozkurt (2014) is used to assess PSTs’ skills. In addition, the LEGO Coding Achievement Assessment (LCAA), developed by the researcher, was administered twice, once as a pre-test before the study and once as a post-test after the study. Quantitative data were analyzed through pre and post-test comparisons. In contrast, qualitative data obtained from audio recordings and interview questions were analyzed and interpreted holistically at the end of the study. Quantitative data served a complementary role in the evaluation of qualitative data.

The timing and sequence of the data collection instruments are presented in Table 1 below.

Table 1. Sequence of Data Collection

| Data Collection Instrument | Before Implementation | During Implementation | After Implementation |
|---|------------------------------|------------------------------|-----------------------------|
| SPST (Scientific Process Skills Test) | X | | |
| SCT (Scientific Creativity Test) | X | | |
| DMST (Decision Making Skills Test) | X | | |
| LCAA (LEGO Coding Achievement Assessment) | X | | X |
| Semi-Structured Questions | | X | |
| Interviews | | | X |

The SPST consists of 25 multiple choice questions, each with four options, and its reliability was calculated with a Cronbach’s alpha coefficient of 0.81.

Each correct response is scored 4 points, while each incorrect response is scored 0 points. The maximum possible score on the test is 100.

The SCT used in this study demonstrated an internal consistency, Cronbach's alpha of 0.89 and a reliability coefficient of 0.87. The test consists of seven open ended questions, and the responses were evaluated based on the percentage of participants who provided each response. For the first three questions, responses given by fewer than 10% of participants received 2 points, while those given by higher percentages received 1 point. In the fourth question, responses with frequencies below 10% were assigned 1 point, whereas more common responses received 0 points. For the fifth question, drawings observed in fewer than 10% of participants received 1 point, while the remaining drawings received 3 points each. In the sixth question, responses occurring in fewer than 10% of participants received 4 points; those in the 10-20% range received 2 points; and those with higher frequencies received 0 points. For the seventh question, each function identified was awarded 3 points. Since the test, which aims to measure creativity, does not specify a maximum number of possible responses, there is no maximum obtainable score.

The DMST used in the study was developed by Bozkurt (2014) and consists of 11 multiple choice questions. Each correct answer is scored as 1 point, while unanswered or incorrectly answered questions are scored as 0 points. No modifications were made to the scoring procedure in this study. The KR-20 internal consistency coefficient of the test is 0.71.

The LCAA used in the study was developed by the researcher and consists of four open ended questions, three of which assess coding skills and one of which addresses science content knowledge. The three coding questions were prepared using screenshots of code that participants had previously learned during their LEGO Mindstorms coding training. The fourth question concerns a specific property of water, one of the fundamental science topics participants would encounter in the study's scenario. The responses obtained from this test were scored as follows: "Correct = 2 points," "Partially Correct = 1 point," and "Incorrect or blank = 0 points." The primary purpose of including codes that participants had previously encountered in the test was to evaluate, through pre and post-test comparisons, the extent to which their knowledge was translated into practice and to determine how the use or non-use of this knowledge contributed to their overall performance.

The researchers developed the scenario used in this study and comprises seven tasks organized into three main headings. Through the tasks presented in the worksheet, PSTs were expected to attribute the main character in the scenario to their robots and to carry out the assigned tasks using their coding knowledge. A LEGO mat to be used alongside the scenario was also designed and provided by the researcher as instructional material. Figures

and materials from the LEGO Mindstorms EV3 Education Core Set and the LEGO Education Teacher’s Kit were utilized to represent the scenario. The semi structured questions included prompts on the difficulties PSTs encountered while performing the tasks, the solutions they generated, the commands they used, situations in which they disagreed with their teammates when deciding on commands, and how they resolved such disagreements. Within the scope of the tasks, it was aimed to measure and evaluate participants’ basic coding skills, engineering skills, awareness of water resources, wastewater, and the water cycle, as well as their ability, as future science teachers, to propose solutions to resource-related problems and to demonstrate interdisciplinary approaches between science and engineering fields.

The data consisted of recordings of approximately two hours each. These recordings were analyzed from several perspectives. First, PSTs’ views on the scenario were examined. Their opinions on science teaching through scenario based instruction were also considered. In addition, their views on integrating LEGO robotics applications into the STEM curriculum were analyzed. The analysis also focused on the difficulties they encountered during coding. Their decision making processes were examined as well. Situations in which they disagreed with their teammates were identified. How they resolved these conflicts was also explored. Furthermore, the solution oriented suggestions they proposed for each task were analyzed. The alternative solutions they pursued when their initial ideas failed were also considered. Finally, observations of their problem solving skills and scientific process skills were included in the analysis. Table 2 below presents the data collection instruments used in the study.

Table 2. Data Collection Instruments and Their Purposes

| Data Type | Data Collection Instrument | Purpose |
|--------------------------|---|---|
| Quantitative Data | Scientific Process Skills Test (SPST) | To measure PSTs’ scientific process skill levels. |
| | Scientific Creativity Test (SCT) | To assess PSTs’ scientific creativity and imagination, which constitute one of the core components of the study. |
| | Decision-Making Skills Test (DMST) | To evaluate the decision making skills that PSTs employ within the context of the scenario. |
| Qualitative Data | LEGO Coding Achievement Assessment (LCAA) Pre-test and Post-test | To examine whether there is an improvement in coding achievement as PSTs’ initial knowledge evolves and becomes more applicable through experience and practice. |
| | Scenario, Semi-Structured Interview Questions, and Audio Recordings | To gain a deeper understanding of PSTs’ experiences during the implementation process, including the challenges they encountered, the inconsistencies they faced, and the strategies they followed to overcome these challenges. Additionally, to analyze their decision making processes through their reflective responses, verbal expressions, and problem-solving approaches. |

Findings

Quantitative Findings

The study's quantitative findings were examined in terms of PSTs' scientific process skills, scientific creativity, and decision-making abilities. These three variables were analyzed to better understand their potential relationship with PSTs' performance in a scenario based, LEGO robotics supported learning environment.

SPST Results

An examination of the scientific process skills indicates that PSTs generally demonstrated moderate to high levels of performance (Table 3). The mean test score was calculated as 76.6, and more than half of the students scored above the arithmetic mean. This finding suggests that the PSTs had a relatively strong baseline in fundamental scientific process skills.

Table 3. Scientific Process Skills Test Results

| PST | Score | PST | Score |
|-------|-------|--------|-------|
| PST1 | 80 | PST 7 | 60 |
| PST 2 | 76 | PST 8 | 80 |
| PST 3 | 88 | PST 9 | 84 |
| PST 4 | 80 | PST 10 | 84 |
| PST 5 | 96 | PST 11 | 56 |
| PST 6 | 60 | PST 12 | 76 |

SCT Results

The results of the SCT show a different pattern than those of the SPST. According to the findings, the mean score was 33.25, and more than half of the participants scored below the average (Table 4). Notably, only four PST performed above the mean, suggesting that scientific creativity was relatively limited across the group. A closer examination reveals that participants frequently provided more routine, commonly expected responses on several items, whereas creative and original answers were comparatively less frequent. However, the presence of some innovative ideas in specific tasks, such as bicycle designs incorporating energy transformation mechanisms or alternative experimental testing approaches, indicates that participants' creative potential was not entirely low, but rather unevenly distributed across the group.

Table 4. Scientific Creativity Test Results

| PST | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Total |
|--------|----|----|----|----|----|----|----|-------|
| PST1 | 0 | 2 | 1 | 2 | 4 | 0 | 3 | 12 |
| PST 2 | 5 | 6 | 3 | 6 | 3 | 0 | 9 | 32 |
| PST 3 | 3 | 2 | 7 | 1 | 16 | 4 | 21 | 54 |
| PST 4 | 15 | 15 | 9 | 6 | 7 | 0 | 15 | 67 |
| PST 5 | 5 | 6 | 3 | 2 | 26 | 0 | 12 | 54 |
| PST 6 | 5 | 3 | 0 | 4 | 20 | 0 | 3 | 35 |
| PST 7 | 4 | 3 | 4 | 3 | 4 | 0 | 3 | 21 |
| PST 8 | 4 | 4 | 2 | 0 | 4 | 2 | 3 | 19 |
| PST 9 | 1 | 5 | 3 | 1 | 2 | 8 | 6 | 26 |
| PST 10 | 3 | 6 | 5 | 3 | 4 | 0 | 6 | 27 |
| PST 11 | 3 | 7 | 4 | 4 | 4 | 0 | 3 | 25 |
| PST 12 | 0 | 4 | 4 | 6 | 4 | 0 | 9 | 27 |

An example of a machine designed by the PSTs within the scope of the SCT is presented in Figure 1.

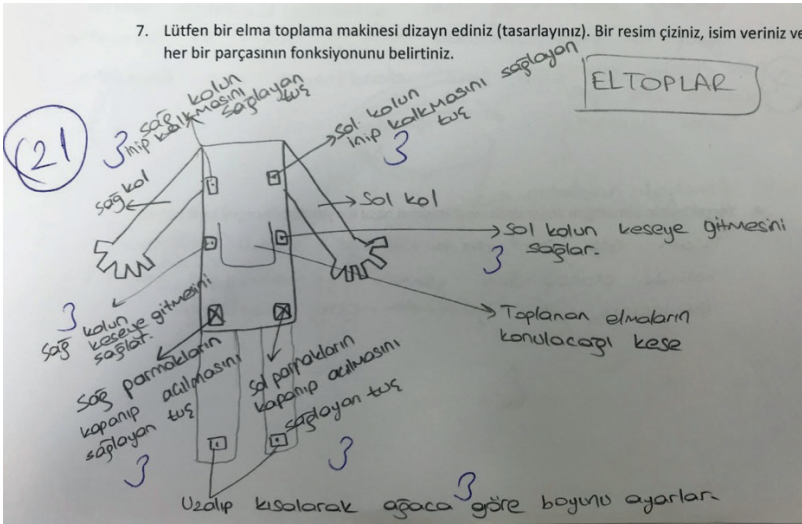


Figure 1. An example of a machine designed by the PST

DMST Results

The findings related to decision making skills indicate a more balanced distribution. The mean number of correct responses was 6.75; six participants scored above the average, while the other six scored below it (Table 5). This result suggests that the group exhibited heterogeneous decision making abilities. Furthermore, item-level analyses revealed that certain questions were particularly challenging for the participants. For instance, “the highest number of incorrect responses was observed in Question 2,” with 10 out of 12 participants answering this item incorrectly. Similarly, the high error rate

observed in Question 7 indicates that participants experienced difficulties, particularly in multi criteria decision making processes.

Table 5. Decision Making Skills Test Results

| PST | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 |
|--------|----|----|----|----|----|----|----|----|----|-----|-----|
| PST1 | T | F | T | T | T | T | F | T | T | T | T |
| PST 2 | T | F | F | T | F | T | F | T | T | T | F |
| PST 3 | T | F | F | T | F | T | T | T | T | T | F |
| PST 4 | T | T | T | T | F | T | T | T | T | T | T |
| PST 5 | T | F | T | F | F | T | F | F | T | T | T |
| PST 6 | F | F | T | F | T | T | F | T | F | F | F |
| PST 7 | T | F | T | F | T | T | F | T | T | T | T |
| PST 8 | T | F | T | T | T | F | F | T | F | T | F |
| PST 9 | T | F | F | T | T | T | T | T | F | T | T |
| PST 10 | T | T | T | F | F | T | F | F | F | F | F |
| PST 11 | T | F | T | T | F | T | F | T | T | F | F |
| PST 12 | T | F | T | T | F | T | T | T | F | T | F |

Overall, it can be concluded that PSTs demonstrated relatively strong scientific process skills, moderate levels of decision making abilities, and comparatively limited scientific creativity. This pattern suggests that the creativity dimension, in particular, needs to be more strongly supported, especially within STEM and design based learning environments.

Qualitative Findings

The qualitative findings of the study are based on the analysis of the LCAA, scenario based implementations, semi-structured interview questions, and audio recordings. These findings reveal both changes in PSTs' coding performance and their experiences within the scenario based robotic learning environment.

LCAA

The findings obtained from the LCAA were analyzed with a focus on the transformation of PSTs' coding understanding throughout the implementation process, rather than solely on correctness levels. At the beginning of the study, PSTs experienced noticeable difficulty in recalling previously learned coding structures. Their responses were generally incomplete and fragmented, often reflecting isolated command-level thinking rather than structured solutions. Participants tended to recall individual commands (e.g., motor activation) but failed to integrate them into a coherent sequence.

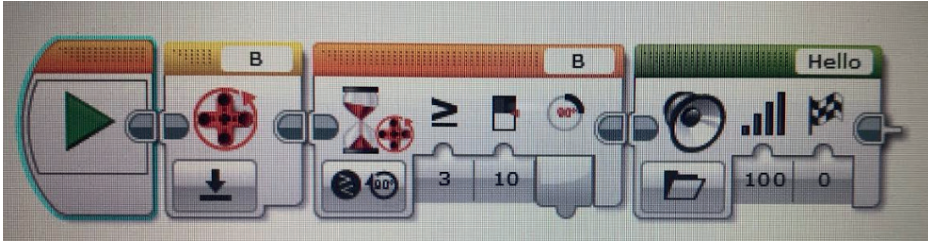


Figure 2. LCAA questions 1

This transition reflects the emergence of procedural engagement, where knowledge is actively reconstructed through interaction with the task and collaboration with peers. By the end of the implementation, PSTs demonstrated a more structured and integrated understanding of coding. Their responses evolved into logically organized sequences, often including multiple steps and coordinated use of commands. Participants were able to align coding structures with task requirements, indicating not only improved recall but also functional understanding.

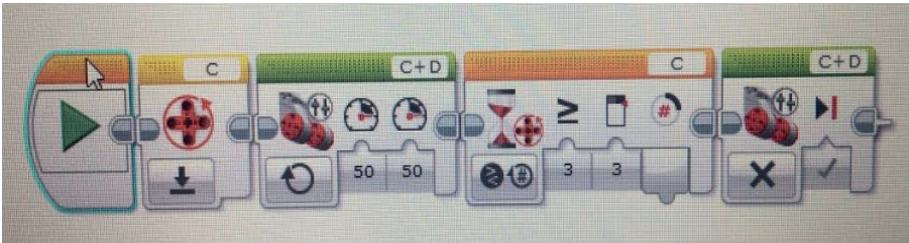


Figure 3. LCAA questions 2

This development suggests a shift from fragmented knowledge toward system-level reasoning, where coding is understood as an interconnected process rather than a set of isolated commands. A similar transformation was observed in the science content-related item of the LCAA. At the beginning of the study, PSTs showed limited understanding of the targeted water-related concept, and many responses included uncertainty or incomplete explanations. However, by the end of the implementation, participants were able to provide clearer, more accurate explanations, often linking the concept to the scenario and to robotic applications.



Figure 4. LCAA questions 3

Although quantitative results (Table 6) indicate a statistically significant increase in scores, the qualitative analysis reveals that this improvement cannot be reduced to numerical gains alone. Instead, it reflects a deeper transformation in PSTs' cognitive processes, including the transition from recall to reasoning, from individual thinking to collaborative problem solving, and from isolated knowledge to integrated understanding.

Table 6. Results of the LCAA Pre-Test and Post-Test

| Test | N | Mean | Std. Deviation | df | p |
|-------------|----------|-------------|-----------------------|-----------|----------|
| Pre-test | 12 | 0.44 | 0.37 | 11 | 0.000 |
| Post-test | 12 | 1.90 | 0.17 | | |

Quantitative analyses also support this improvement: while the pre-test mean score was quite low (0,44), the post-test mean increased significantly (1,9), and this difference was found to be statistically significant (Table 6). This finding suggests that success in robotic coding is directly related to practice and repetition.

Overall, the LCAA findings demonstrate that learning in this context is best understood as a progressive and reconstructive process. The integration of scenario-based learning and LEGO robotics enabled PSTs to activate prior knowledge, restructure their understanding through experience, and develop more coherent and functional coding and conceptual knowledge.

Task Based Scenario Process

In the study, the scenario based text developed by the researchers, along with the semi structured questions it contained, consisted of seven tasks organized under three main headings. Through these tasks, participants were expected to assign the main character in the scenario to their robots and to carry out the assigned tasks using their coding knowledge. A LEGO mat to be used alongside the scenario was also designed by the researchers and provided as instructional material.

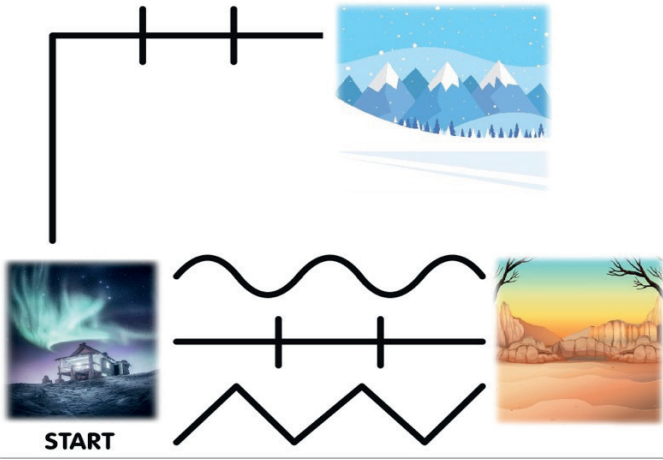


Figure 5. LEGO Mat designed in alignment with the scenario

The tasks carried out within the scope of the scenario revealed PSTs’ problem solving, collaboration, and decision making processes. The majority of participants reported difficulty recalling coding commands and challenges selecting commands and adjusting parameters. However, they also stated that they reached solutions through trial and error and peer collaboration.

For instance, one participant’s statement, “I forget when I do not use the program regularly,” and another’s remark, “I found which one works better by trying,” indicate that learning occurred in an experiential and process oriented manner. The audio recordings also support this finding:

“Let’s try the large motor first... if it doesn’t work, we’ll try the medium motor.”

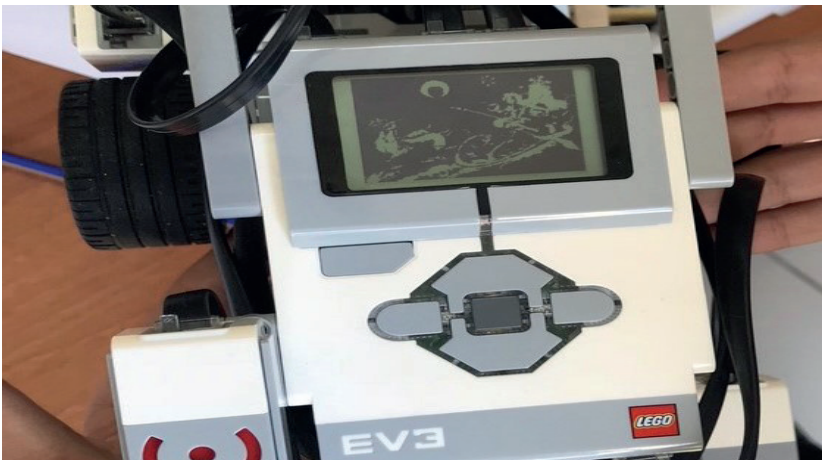


Figure 6. Findings related to the second phase of the implementation

These findings indicate that 21st-century skills such as collaboration, communication, and problem solving were actively utilized.

Abstraction and Symbolization Skills

A significant proportion of participants reported initially thinking concretely. However, they transitioned to abstract thinking throughout the process. In particular, the symbolization of concepts such as states of matter and the formation of rain using LEGO components deepened conceptual understanding.

One participant's statement clearly illustrates this transformation:
"Once we represented water symbolically rather than concretely, everything was resolved."



Figure 7. *Visual of the rain phenomenon symbolized in the fourth phase of the implementation*

Decision Making and Analytical Thinking Process

In Task II, participants were required to select the fastest route among several alternatives, and they were expected to use mathematical reasoning ($X = V \cdot t$) in this process. Although some participants initially chose incorrect paths, the fact that all participants reached the correct decision by the end of the process indicates that analytical thinking developed over time.

Additionally, although disagreements occurred among participants, these were generally resolved through discussion, trial and error, and consensus.

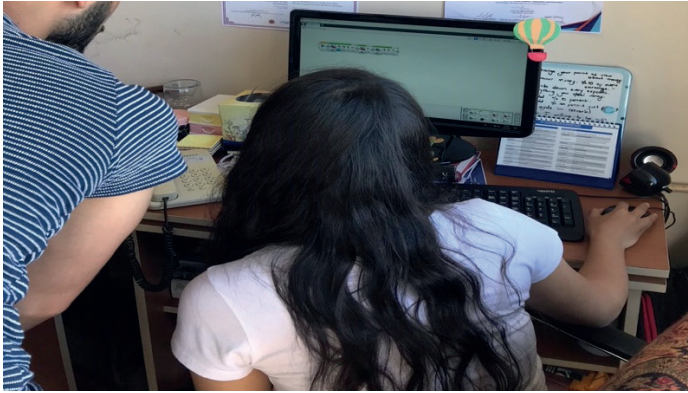


Figure 8. Participants performing coding for light calibration

Science Concepts and Environmental Awareness

It was observed that PSTs initially had limited knowledge regarding water resources, the water cycle, and environmental sustainability. Many participants responded to these questions with “*I don’t know.*”

However, awareness developed throughout the process. For example, one participant stated:

“I had never thought about where water comes from before.”

This finding indicates that the scenario-based approach is effective not only at the cognitive level but also in terms of affective outcomes and awareness.

Engineering Design and STEM Integration

All participants developed water towers and various engineering designs. These designs indicate the development of skills such as environmental problem-solving, systems thinking, and innovative approaches. For example, one participant proposed an advanced solution, stating:

“Water collecting towers from the air... condensation using liquid nitrogen...”

General Participant Views

Participants generally evaluated the implementation as enjoyable, instructive, engaging, and conducive to lasting learning. The following statement is particularly noteworthy:

“Storytelling draws people into the event... it leaves a lasting impression on children.”

PST1: *“I liked it very much. I found it quite creative. Effective and engaging activities were carried out by demonstrating science concepts.”*

PST2: *“I found the scenario nice and enjoyable. Teaching coding through a scenario is a good method both for robotics coding education and for teaching a science topic.”*

PST3: *“I found it enjoyable and instructive. It was an activity that taught through doing and observing.”*

PST4: *“I support the integration of a scientific framework into science education. I think the scenario in this study has a creativity-enhancing and innovative effect, and that learning coding in this way leads to more lasting learning.”*

PST5: *“Storytelling and scenario design draw people into the event. If such a method can make even adults very happy and thoughtful, I believe it is a very effective way to leave a lasting impression on children.”*

PST6: *“Teaching the lesson through this kind of scenario would attract students’ attention, making the lesson more efficient and enjoyable. I would very much like to teach lessons this way in the future. The scenario was excellent.”*

PST9: *“I liked your study and your scenario. I think it was different, creative, and enjoyable.”*

PST10: *“The scenario was well designed. There were interesting and thought-provoking events. It was useful to develop solutions to these situations.”*

PST11: *“I answered the survey questions without getting bored; it was enjoyable. I especially liked the storytelling and design part. Answering the questions in the robotics achievement test helped me recall what I had forgotten.”*

This finding clearly demonstrates the pedagogical effectiveness of scenario based learning.

Overall, the qualitative findings indicate that the scenario- and robotics-supported learning environment enhances not only coding skills but also multidimensional competencies, including problem solving, collaboration, abstract thinking, engineering design, and environmental awareness.

Conclusion

This study aimed to examine the effects of PSTs’ scientific process skills, scientific creativity, and decision-making skills on their coding achievement in a scenario based, LEGO robotics supported science learning environment, and to explore their views of this learning environment.

When the findings are evaluated overall, it is observed that PSTs possess adequate levels of scientific process and decision-making skills, whereas their scientific creativity appears relatively limited. However, it was also determined that during the implementation process, PSTs successfully performed tasks requiring scientific creativity. This indicates that ER supported learning

environments not only reveal existing skills but also foster the development of creativity and design skills.

Findings related to scientific process skills indicate that PSTs effectively utilized skills such as observation, measurement, data collection, hypothesis formation, and experimental procedures throughout the implementation. This result partially differs from studies in the literature suggesting that LEGO has limited effects on cognitive development (Hussain, Lindh, & Shukur, 2006); instead, it demonstrates that individuals with strong scientific process skills perform more successfully in robotics environments. Similarly, the study conducted by Okkesim (2014), which highlights that robotics supported learning enhances scientific process skills, supports these findings.

When findings on scientific creativity are examined, it is evident that PSTs initially performed poorly. However, their successful performance in STEM-based design tasks throughout the process indicates that scientific creativity can be developed through appropriate learning environments. This finding is consistent with studies emphasizing the positive relationship between scientific process skills and scientific creativity (Roberts, 2003; Aktamış & Ergin, 2007). Therefore, it can be concluded that both suitable learning environments and instructional approaches that support process skills should be considered together to foster scientific creativity.

In terms of decision making skills, the findings indicate that PSTs effectively used them, particularly during the implementation process. It was determined that group work and teacher support played significant roles in decision making. This aligns with studies emphasizing the influential role of the teacher in the learning process (Pedersen, 1998, as cited in Hussain, Lindh, & Shukur, 2006). Furthermore, participants' tendency to adopt a collaborative and consensus-oriented approach during problem solving suggests that robotics learning environments promote cooperation rather than competition.

One important finding of the study is that PSTs tend to forget their coding knowledge when they do not engage in practice; however, they can regain this knowledge after a short but intensive implementation process. This highlights the critical role of practice and experience based learning in ensuring retention. Additionally, it was observed that robotics supported learning enhances not only technical skills but also conceptual understanding. In particular, the fact that all participants were able to correctly explain concepts such as "*the specific property of water*" at the end of the implementation clearly demonstrates the effectiveness of this approach in science education.

However, it was also found that PSTs initially lacked sufficient knowledge regarding fundamental science topics such as water resources, the water cycle, and environmental issues. This finding underscores the need to strengthen content knowledge in science education.

When PSTs' views on the scenario based robotics learning environment were examined, it was found that they perceived the process as enjoyable, motivating, and instructive. Contrary to studies suggesting that competition increases motivation (Giannakopoulos, 2009), this study revealed that participants' motivation was driven by collaboration rather than competition. Statements such as “*we made decisions together*” and “*we found a common solution*” indicate that such environments support collaboration, communication, and social interaction within the framework of 21st-century skills (Çepni & Ormancı, 2018).

In conclusion, this study demonstrates that scenario based, robotics supported science learning environments provide a powerful pedagogical approach that enables the effective use of scientific process skills, fosters scientific creativity, supports decision making, and strengthens conceptual learning. At the same time, these environments allow students to reconstruct their prior knowledge (Altın & Pedaste, 2013).

Recommendations

In conclusion, in line with the findings obtained from the study, the following recommendations are proposed:

Conducting similar studies with students and different age groups,

Promoting the use of LEGO robotics applications by both students and teachers,

Placing greater emphasis on studies that enhance scientific creativity for effective science education,

Expanding educational efforts to train more knowledgeable teachers on topics such as “water” and “recycling” within the science curriculum.

References

- Aktamış, H. & Ergin, Ö. (2007). Bilimsel süreç becerileri ile bilimsel yaratıcılık arasındaki ilişkinin belirlenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33(33), 11-23.
- Altın, H. & Pedaste, M. (2013). Learning approaches to applying robotics in science education. *Journal of Baltic Science Education*, 12(3), 365-377.
- Aydın, M., (2018). Lego Robotik Uygulamaları ile STEM Eğitimi. S. Çepni (Ed.), *Kuramdan Uygulamaya STEM Eğitimi* (ss. 369-388). Ankara: Pegem Akademi.
- Barreto, F. & Benitti, V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education* 58(3), 978-988.
- Bilotta, E., Gabriele, L., Servidio, R. & Tavernise, A. (2009). Edutainment robotics as learning tool. M. Chang (Ed.), *Transactions on edutainment III* (pp. 25-35). Berlin, Heidelberg: Springer.
- Bozkurt, E. (2014). Mühendislik tasarım temelli fen eğitiminin fen bilgisi öğretmen adaylarının karar verme becerisi, bilimsel süreç becerileri ve sürece yönelik algılarına etkisi (Unpublished doctoral dissertation), Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Burns, J. C., Okey, J. R. & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, 22(2), 169-177.
- Cejka, E., Rogers, C. & Portsmore, M. (2006). Kindergarten robotics: Using robotics to motivate math, science, and engineering literacy in elementary school. *International Journal of Engineering Education*, 22(4), 711.
- Çepni, S. (2018). *Kuramdan Uygulamaya STEM Eğitimi*. Pegem Akademi: Ankara.
- Geban, Ö., Aşkar, P. & Özkan, İ. (1992). Effects of computer simulations and problem-solving approaches on high school students. *The Journal of Educational Research*, 86(1), 5-10.
- Giannakopoulos, N. (2009, September). Experiences from WRO 2009 competition and verifications about the robotics incorporation in the school. *Book of Abstracts of the Conference "Lessons Learnt from the Terecop Project and New Pathways into Educational Robotics across Europe*. Athens, Greece.
- Hu, W. & Adey, P. (2002). A scientific creativity test for secondary school students. *International Journal of Science Education*, 24(4), 389-403.
- Hussain, S., Lindh, J. & Shukur, G. (2006). The effect of LEGO training on pupils' school performance in mathematics, problem solving ability and attitude: Swedish data. *Journal of Educational Technology & Society*, 9(3), 182-194.
- Johnson, J. (2003). Children, robotics, and education. *Artificial Life and Robotics*, 7(1-2), 16-21.
- Jung, S. (2013). Experiences in developing an experimental robotics course program for undergraduate education. *IEEE Transactions on Education*, 56(1), 129-136.

- Karp, T., Gale, R., Lowe, L. A., Medina, V. & Beutlich, E. (2010). Generation NXT: Building young engineers with LEGOs. *IEEE Transactions on Education*, 53(1), 80-87.
- Klassner, F. & Anderson, S. D. (2003). Lego MindStorms: Not just for K-12 anymore. *IEEE Robotics & Automation Magazine*, 10(2), 12-18.
- Koç Şenol, A. & Büyük, U. (2015). Robotik destekli fen ve teknoloji laboratuvar uygulamaları: Robolab. *Turkish Studies*, 10(3), 213-236.
- Okkesim, B. (2014). Fen ve teknoloji öğretiminde robotik uygulamaları (Unpublished master's thesis), Erciyes Üniversitesi Eğitim Bilimleri Enstitüsü, Kayseri.
- Papert, S. (1993). *Mindstorms: Children, computers, and powerful ideas* (2nd ed.). New York: Basic Books.
- Roberts, L. (2003). Creativity, *Tech Directions*, 63 (3).
- Silik, Y. (2016). Eğitsel Robotik Uygulamalarının Fen Bilgisi Öğretmen Adaylarının Problem Çözme Becerilerine Etkisi (Unpublished master's thesis), Karadeniz Teknik Üniversitesi, Trabzon.
- Williams, D. C., Ma, Y., Prejean, L., Ford, M. J., & Lai, G. (2007). Acquisition of physics content knowledge and scientific inquiry skills in a robotics summer camp. *Journal of Research on Technology in Education*, 40(2), 201-216.
- Wood, S. (2003). *Robotics in the classroom: A teaching tool for K-12 educators*, Symposium of Growing up with Science and Technology in the 21st Century, Virginia.
- Yıldırım, A. & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.