Journal of Pedagogical Research
Volume 10, Issue 5, 2026
https://doi.org/10.33902/JPR.202648644

Research Article

Fostering scientific inquiry and curiosity through story-based STEM activities in preschool education: A design-based approach

Merve Şahin

Çanakkale Onsekiz Mart University, Bayramiç Vocational School, Department of Child Care and Youth Services, Child Development Program, Türkiye (ORCID: 0000-0001-9394-9511)

Narrative contexts are increasingly used to frame early science, technology, engineering, and mathematics (STEM) learning, yet little is known about how children of different ages respond to the same story world. This design-based study examined the scientific process skills, problem-solving strategies, and curiosity–exploration behaviors of 30 children aged 3–6 years, in three age groups, during four weeks of STEM activities built around *Alice's Adventures in Wonderland*, together with the views of 15 preschool teachers. A seven-dimension observation form, an activity evaluation form, semi-structured teacher interviews, field notes, and children's products were analyzed descriptively and thematically and integrated in a joint display. Observation scores increased in all three age groups. The largest gain appeared among 4–5-year-olds, whereas most 5–6-year-olds reached the high-performance band, suggesting a ceiling tendency. Across sources, children moved from handling materials by their surface features toward predicting, testing, redesigning, and explaining their solutions, and their curiosity became more analytical with age. Teachers judged the activities feasible but asked for ready-to-use plans, age-appropriate materials, and visual supports. The study proposes five design principles for age-sensitive story-based STEM. Because no control group was used, these within-group patterns are descriptive and cannot be attributed to the activities alone.

Keywords: Alice's Adventures in Wonderland; Design-based research; Preschool education; Scientific inquiry skills; Story-based STEM

Article History: Submitted 21 August 2026; Revised 10 October 2026; Accepted 30 October 2026; Published online 5 November 2026

1. Introduction

Early childhood is a period in which the foundations of cognitive, linguistic, and motor development are laid down quickly, and dispositions such as curiosity, exploration, questioning, experimentation, and problem solving begin to take shape. The experiences offered to children in these years influence how they make sense of the world, notice relationships among events, try out materials, and generate solutions to the problems they meet. It is therefore not surprising that education in science, technology, engineering, and mathematics (STEM) has become a prominent strand of early childhood research (Bers, 2021; Larkin & Lowrie, 2022; McClure et al., 2017; Wan et al., 2021).

Address of Corresponding Author

Merve Şahin, PhD, Çanakkale Onsekiz Mart University, Bayramiç Vocational School, Department of Child Care and Youth Services, Child Development Program, Bayramiç, 17700 Çanakkale, Türkiye.

✉ mervesahin0617@gmail.com

How to cite: Şahin, M. (2026). Fostering scientific inquiry and curiosity through story-based STEM activities in preschool education: A design-based approach. *Journal of Pedagogical Research, 10*(5), 93–120. https://doi.org/10.33902/JPR.202648644

Copyright © 2026 by the author(s). This is an **open-access** article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Rather than transmitting abstract scientific concepts, preschool STEM education invites children to handle concrete materials, observe, predict, experiment, and develop possible solutions to the problems in front of them. Children in this age range already show early indicators of scientific thinking: they examine objects, connect causes with effects, compare materials, watch what happens when they try something, and explain the outcome in their own words (Eshach & Fried, 2005; Gelman & Brenneman, 2004; Trundle & Smith, 2017). Seen this way, early STEM cannot be reduced to science or mathematics content; it is a holistic approach that supports process-oriented skills such as inquiry, curiosity, exploration, creativity, and problem solving.

Scientific inquiry sits at the center of preschool STEM practice. Observing, asking questions, predicting, experimenting, comparing results, and offering explanations all play a part in children's efforts to understand their surroundings (Gelman & Brenneman, 2004; Ramanathan et al., 2022). For these skills to grow, however, children need concrete, play-based, and meaningful contexts suited to their developmental level. Cremin et al. (2015) emphasize that scientific learning in the early years is supported by play, creativity, dialogue, and inquiry. Curiosity and wonder are likewise recognized as powerful sources of scientific learning, although further empirical work is needed to show how they can be sustained pedagogically (Bjerknes et al., 2024).

From a developmental-psychological perspective, these behaviors are early forms of scientific reasoning, in which children begin to coordinate their ideas with evidence (Zimmerman, 2007), and they are energized by curiosity, which Loewenstein (1994) describes as a motivational response to a perceived gap in knowledge. Narrative may matter here for cognitive as well as motivational reasons. Bruner (1986) argued that narrative is one of two fundamental modes of thought through which people organize experience; a story problem can therefore give young children a ready-made structure of goals, obstacles, and consequences within which to place their own trials. Adult guidance inside such a structure resembles the scaffolding described by Wood et al. (1976), in which the adult manages the parts of a task that lie beyond the child's current reach so that the child can concentrate on the parts within it.

Problem solving is also a core component of early STEM experiences. Faced with a real or fictional problem, preschool children may reorganize materials, try alternative strategies, modify a solution that does not work, and rebuild their product. Studies of design thinking and STEM-based practice in preschool settings indicate that such activities can support early problem-solving, creative-thinking, and engineering-design skills (Akcay Malcok & Ceylan, 2022; Yalçın & Erden, 2021). Early STEM is not a purely cognitive undertaking, either. In an integrated STEM and social–emotional learning program with children aged 48–60 months, teachers linked children's rising observation scores to collaboration, empathy, and perseverance, and to emotionally responsive teaching such as emotional coaching (Şahin, 2026a). Well-structured STEM experiences thus do more than build domain-specific knowledge; they make children's thinking processes visible, including observation, prediction, inference, comparison, design development, and explanation of results.

Story-based learning offers a promising entry point for this kind of work. Stories help young children sustain attention, concretize abstract ideas, make problem situations meaningful, and frame the learning process (Cremin et al., 2017). A narrative structure allows a child to engage with a task not simply to finish it, but to overcome an obstacle, help a character, or make sense of events in the story. Marcus et al. (2021) showed that narrative reflection on a problem-solving experience can increase children's STEM-related talk both during the activity and in later recall. A story, in other words, may be more than an attention-getter; it may provide a functional context for making sense of, remembering, and re-expressing STEM content.

One appealing feature of story-based science learning is that it invites children to question the relationship between fiction and the real world. Studies that set extraordinary events in children's books against real-world experiments suggest that unusual situations in stories can become starting points for scientific inquiry (Blanquet & Picholle, 2012; Howitt et al., 2011). Integrating story elements into problem-solving activities can also support STEM talk, recall, and conceptual

connection-making (Altun, 2019; Marcus et al., 2021; Miller-Goldwater et al., 2024). Work on digital storytelling, robotic storytelling, and story-based programming points in the same direction: narrative approaches situate children's scientific thinking, computational thinking, task planning, and solution generation within a meaningful context (Cheng & Chuang, 2019; Esther-del-Moral-Pérez et al., 2026; Tengler et al., 2022; Yang et al., 2023).

Existing studies report that narrative, digital storytelling, and gamified approaches can yield supportive outcomes for preschool children's learning interest, scientific concept development, and problem solving (Eliza et al., 2025; Shemy, 2020; Vaknin-Nusbaum & Nevo, 2017). Digital storytelling has been associated with preschool children's motivation and achievement in learning scientific concepts, and picture-based science story applications with science process skills (Eliza et al., 2025; Shemy, 2020). A recent study in the Turkish context added that, for story-based and digitally supported science, technology, engineering, arts, and mathematics (STEAM) practice to work, pre-service teachers' positive attitudes toward technology are not enough; their domain-specific beliefs about the pedagogical value of children's literature for mathematics and STEAM learning play a decisive role in their digital competence (Şahin, 2026b). The quality of story-based STEM/STEAM practice, then, depends not only on children's engagement with narrative but also on educators' capacity to turn stories into developmentally appropriate learning experiences. Earlier work with child development students points the same way: they found STEM activities hard to plan until they had acquired the relevant content knowledge, after which they felt able to implement them (Şahin, 2021). Even so, much of the existing work focuses on specific learning outcomes or on general interest in activities. Multi-source evidence on how story-based STEM practice produces different patterns of scientific process skills, problem solving, and curiosity across age groups remains limited (Zhou & Yang, 2025).

Developmental differentiation across age groups deserves particular attention when story-based STEM practice is designed. Concrete materials, brief instructions, and direct manipulation may be more decisive for children aged 3–4, whereas symbolic play, metaphorical thinking, classification, and comparison may become more visible at 4–5 years (Cremin et al., 2015; Gelman & Brenneman, 2004; Trundle & Smith, 2017). By 5–6 years, children may establish more complex cause–effect relationships, justify their designs, test their products, and revise solutions that do not work (Akcay Malcok & Ceylan, 2022; Vidal Carulla et al., 2021; Yalçın & Erden, 2021). In story-based and gamified STEM studies, age is likewise treated as an important variable in relation to symbolic representation, task structure, and cognitive flexibility (Esther-del-Moral-Pérez et al., 2026; Howitt et al., 2011; Yang et al., 2023; Zhou & Yang, 2025). These considerations suggest that story-based STEM activities should be organized not as uniform designs but as age-sensitive, pedagogically differentiated arrangements.

A one- or two-year difference is substantial at these ages. Between three and six years, executive functions such as working memory, inhibitory control, and cognitive flexibility develop rapidly (Diamond, 2013). This affects how many steps of an instruction a child can hold in mind, how long the child persists after a failed attempt, and whether the child can switch to an alternative strategy. Language growth over the same period changes how fully children can put predictions and explanations into words, which matters because explanation is one of the main windows onto their reasoning (Gelman & Brenneman, 2004). The ability to coordinate a claim with evidence is also still emerging and becomes more explicit across early and middle childhood (Zimmerman, 2007). For these reasons, the three age groups in the present study did not receive the same activities at different levels of difficulty; activity content, materials, instructions, and session length were differentiated for each group (see Story-Based STEM Activities), and age-group results are compared descriptively rather than as a controlled contrast of age.

Classic children's literature offers a distinctive opportunity in this respect. Lewis Carroll's *Alice's Adventures in Wonderland* provides a rich narrative world that children can readily enter through imagination. Alice's growing and shrinking open natural pathways to size, proportion, and measurement; the small door and key support problem solving and cause–effect reasoning;

the Mad Hatter's tea party can introduce time, sequencing, and order; and mirror and light scenes lead toward reflection, shadow, and optics. Studies that combine all three elements, systematically turning a classic work into STEM activities, adapting those activities for different age groups, and examining them through multiple data sources, are scarce; the present study was designed to bring these elements together.

Choosing a fantasy text for science learning in the early years nevertheless needs justification, because science content is more often introduced to young children through realistic stories. Research on learning from fiction explains why. Preschool children transfer problem solutions from storybooks to the real world less readily when the story is fantastical than when it is realistic (Richert et al., 2009), and anthropomorphic picture books can leave children with less accurate factual knowledge than realistic ones (Ganea et al., 2014). At the same time, children become increasingly sensitive to the causal structure of fictional worlds and use it to judge whether story events could happen in reality (Walker et al., 2015), and reviews of this work conclude that what children take from fiction depends largely on how adults frame the relationship between the story and the real world (Hopkins & Weisberg, 2017). The present design builds on this distinction. *Alice's Adventures in Wonderland* was not used as a source of scientific information. Its fantastical events, such as a door that is too small, a body that changes size, or a world reversed in a mirror, served only as problem frames, and every question raised by the story was answered by testing real materials. Fantasy, in other words, created the puzzling situation that prompted a question, while the materials supplied the evidence. The book also suits young children for practical reasons: its episodes are short and can be presented separately, its central problems involve concrete physical quantities such as size, fit, time, light, and reflection, and it is available in many picture-book adaptations for young readers.

In Türkiye, academic interest in early STEM practice has grown, but more empirical evidence is needed on how such practice can be structured sustainably in classrooms, how teachers evaluate it, and how it can be adapted for different ages (Uğraş & Genç, 2018; Uğraş et al., 2025). The national preschool program, the Türkiye Century Maarif Model (TYMM), emphasizes child-centered, play-based learning that supports holistic development and connects with everyday experience, which offers a strong foundation for linking STEM with stories, play, art, language, science, and mathematics (Ministry of National Education [MoNE], 2024). A large share of existing studies, however, concentrate on teacher views, STEM awareness, teacher self-efficacy, or general implementation opportunities. Far fewer examine children's observable patterns of scientific process skills, problem solving, and curiosity–exploration during integrated classroom STEM experiences built on a classic work of children's literature and compared across age groups, or document how such activities correspond to the outcomes of the national program (see Appendix 1).

The present study addresses this gap by examining how story-based STEM activities designed around *Alice's Adventures in Wonderland* were reflected in the scientific process skills, problem-solving strategies, curiosity and exploration behaviors, and age-related responses of children aged 3–6 years. Using a design-based research approach, it drew on child observation forms, an activity evaluation form, semi-structured interviews with 15 teachers, field notes, and children's products. The study has three intended contributions: an age-sensitive design model that integrates classic children's literature with preschool STEM; a comparative description of how children aged 3–4, 4–5, and 5–6 years responded to the same narrative framework; and a multi-source account in which child observation data, activity evaluation records, teacher interviews, field notes, and children's products are read together. Because the design is descriptive and has no comparison group, no directional hypotheses were tested; instead, the study was guided by five research questions (RQs):

RQ 1) What patterns in children's scientific process skills were observed across the four story-based STEM activities?

RQ 2) What changes in children's problem-solving strategies became visible during the activities?

RQ 3) How did children in the three age groups respond to the same narrative framework?

RQ 4) Through which observable behaviors did curiosity and exploration appear during the activities?

RQ 5) How did teachers evaluate the classroom applicability of the story-based STEM activities?

2. Method

2.1. Research Design

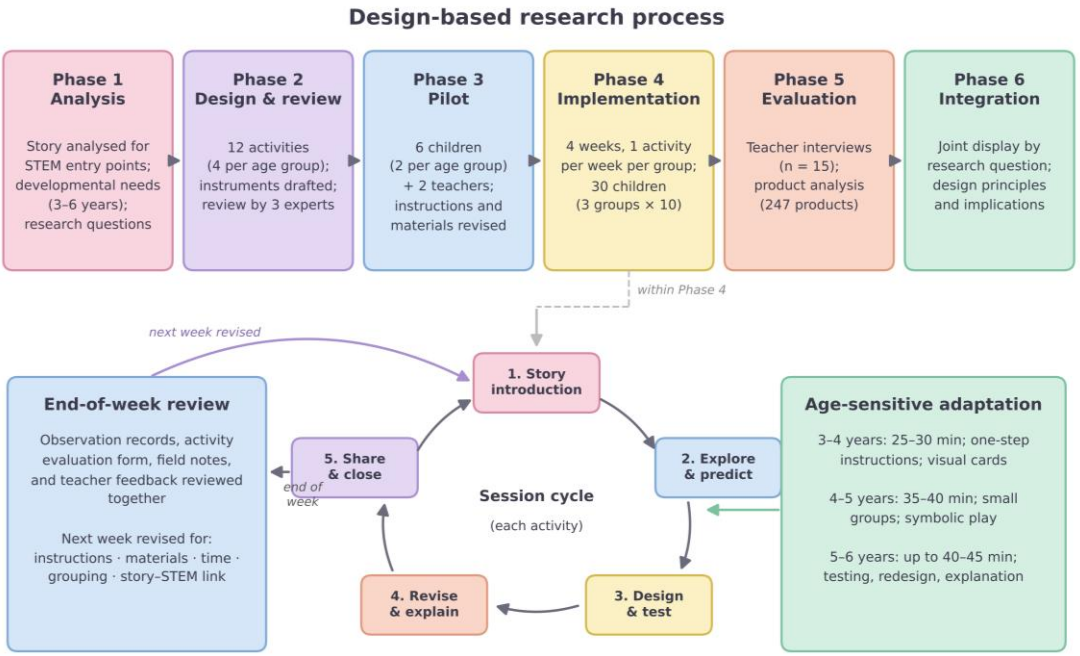
The study followed a design-based research approach to examine how story-based STEM activities worked in a natural preschool classroom, how children of different ages responded to them, and how teachers evaluated the process. Design-based research is a cyclical approach in which educational interventions are designed, implemented, evaluated, and revised in authentic settings on the basis of feedback gathered during implementation (Anderson & Shattuck, 2012; Barab & Squire, 2004).

This design was chosen to examine, in context, how activities developed around *Alice's Adventures in Wonderland* functioned within the ordinary flow of the preschool day and which patterns of scientific process, problem solving, curiosity, and collaboration they made visible among children of different ages. The purpose was not to demonstrate a causal effect of the intervention experimentally but to describe and interpret the developmental and pedagogical patterns that emerged, using multiple data sources. No control group was therefore used; instead, the findings were evaluated through triangulation across child observation forms, the activity evaluation form, teacher interviews, field notes, and children's products (Patton, 2015).

To make the cyclical character of the design visible, child observation records, activity evaluation forms, researcher field notes, and teacher feedback were reviewed together at the end of each implementation week. The following week's activities were then revised in terms of clarity of instructions, choice of materials, time management, group organization, and the link between story and STEM task. Figure 1 shows how these weekly cycles sit within the six phases of the study.

Figure 1

Phases of the design-based research process, the session cycle repeated in each activity, and the weekly review that fed revisions into the next week



2.2. Participants

2.2.1. Child participants

Thirty children attending a summer school program at a preschool in the Marmara Region of Türkiye took part. They were divided into three age groups (3–4, 4–5, and 5–6 years) of 10 children each. Table 1 shows the distribution by age group, sex, and mean age.

Table 1
Distribution of child participants by age group and sex

Age Group	*n*	*Girls*	*Boys*	*Mean Age (Months)*
3–4 years	10	5	5	42.3
4–5 years	10	6	4	54.7
5–6 years	10	4	6	66.5
Total	30	15	15	—

Participants were selected through criterion sampling (Patton, 2015). Inclusion criteria were being between 3 and 6 years of age, attending the summer program regularly, having written parental consent, and being willing to take part. Because the activities and observation instruments in this first design cycle had not yet been adapted for individualized support, children receiving special education services were not included in this cycle; adapting the design for inclusive classrooms is identified as a next step in the Recommendations. No child withdrew. Sex was recorded for descriptive purposes only; no sex differences were hypothesized and, given subgroup sizes of four to six children, no sex-disaggregated analyses were carried out.

The study took place during the preschool's full-day summer program. The story-based STEM sessions were held once a week during the morning activity period, with the classroom teacher present; for the rest of the day, and on the remaining days of the week, children followed the school's usual summer routine. Regular attendance was an inclusion criterion, and all children were present at the first and fourth activities, on which the change scores are based.

2.2.2. Teacher participants

Fifteen preschool teachers took part, selected through criterion sampling (Patton, 2015). Three were the classroom teachers working in the rooms where the activities took place. The other 12 met the criteria of currently working as preschool teachers and agreeing voluntarily to take part, and were invited to the interview process to evaluate the classroom applicability, age-group adaptability, and pedagogical requirements of the activities. This arrangement meant that teacher data were not confined to the implementing teachers' observations but drew on a wider range of professional experience. Teachers were anonymized as T1–T15, and the relevant code is given after each teacher statement in the Findings (see Table 2).

Table 2
Information on the teacher participant group

Variable	*Description*
Number of teachers	15
Coding	T1–T15
Field	Preschool education
Role in the implementation	3 teachers served as classroom implementers; 12 teachers took part in the interview and evaluation process
Data sources	Semi-structured interviews, activity evaluation observation form, and observations of the implementation process
Interview focus	Scientific process skills, curiosity, problem solving, differentiation across age groups, classroom applicability, and curriculum integration

2.3. Story-Based STEM Activities

The implementation lasted four weeks. The intervention consisted of story-based STEM activities adapted from Lewis Carroll's Alice's Adventures in Wonderland. Four activities were planned for each age group, giving 12 activities in total (Table 3). All activities were designed with preschool children's developmental characteristics in mind and included concrete materials, play-based problem situations, short retold story episodes, a prediction–testing–explanation cycle, and group sharing.

Table 3
Story-based STEM activities implemented by age group

Age Group	*Week*	*Activity*	*Story Connection*	*STEM Focus*
3–4 years	1	Growing Chairs	Alice's growing/shrinking scene	Mathematics: size
	2	Magical Color Bottles	The drink bottle scene	Science: color mixing
	3	Lost Keys	The small door and key scene	Science: magnetism
	4	Towers That Do Not Fall	The Queen's palace	Engineering: balance
4–5 years	1	Tea Time in the Time Tunnel	The Mad Hatter's tea party	Mathematics: time
	2	Entrance to the Mysterious Garden	The Wonderland garden	Science: plant growth
	3	Opening and Closing Doors	The small door	Engineering: simple mechanisms
	4	Colorful Journey	Alice's journey	Technology: coding logic
5–6 years	1	The Invisible Path	The mirror world	Science: light and reflection
	2	The Mad Hatter's Bridge	The Hatter's table	Engineering: structure and load-bearing
	3	The Rabbit Lost in the Dark	The White Rabbit	Science: shadow and light
	4	The World Turned Upside Down	The mirror passage	Science: optics

Each activity unfolded in three stages (shown in more detail as a five-step session cycle in Figure 1). First, the relevant episode was retold and the STEM problem was framed within the story. Second, children worked with materials, made predictions, carried out trials, and developed solution strategies. Third, products, observations, and inferences were shared with the group.

The book was not read to the children in its original form, since the full text is long and linguistically demanding for preschool children. Instead, the researcher retold each selected episode in simplified language with visual support, and the story was adapted in two further ways that followed directly from the design. First, only one episode was used in each session, reduced to the event that set up the problem, such as Alice being too large to pass through the little door. Second, the problem posed by the episode was restated as a concrete question that children could answer by testing materials, so that answers came from children's own trials rather than from the story. The closing phase then gave children the opportunity to share and explain what had happened in their trials.

Activities took place once a week, and their length, instructions, materials, and inquiry emphasis were adjusted to the age group (Table 4). Sessions with 3–4-year-olds were the shortest and relied on one-step instructions with visual support, whereas sessions with 5–6-year-olds were the longest and left more time for testing designs, improving products, and explaining results.

Table 4
Age-related adaptations of the story-based STEM activities

Feature	*3–4 Years*	*4–5 Years*	*5–6 Years*
Session length	About 25–30 min	About 35–40 min	About 40–45 min in some sessions
Instructions	Short, concrete, one step at a time, with visual cards	Brief, with modeling and small-group explanation	Multi-step, with open questions about testing and improving
Materials	Large, tactile, easily manipulated pieces	Multipurpose materials that invite symbolic play	Materials that allow testing, strengthening, and redesign
Main play format	Direct manipulation and exploration	Small-group work and symbolic play	Design challenge with testing
Inquiry emphasis	Observing, comparing, simple prediction	Trying alternatives, comparing solution paths	Explaining outcomes and justifying design decisions

Note. Adaptations reflect the activity plans and the revisions made after the pilot implementation (see Pilot Implementation and Instrument Development).

Appendix 1 (Table A1) maps each of the 12 activities onto the components of the TYMM preschool program (MoNE, 2024): field skills, social–emotional learning skills, and dispositions. The mapping shows that the activities addressed program outcomes in mathematics, science, Turkish language, movement and health, art, and social–emotional learning within existing activity time rather than adding content outside the program.

2.4. Data Collection

Data were collected through five sources: the STEM–Scientific Inquiry Observation Form, the Activity Evaluation Observation Form, the Semi-Structured Teacher Interview Form, field notes, and children's products.

2.4.1. STEM–Scientific Inquiry Observation Form

This form was developed to assess children's scientific inquiry and problem-solving performance during the activities. It covers seven dimensions: observation, prediction, inference, comparison, measurement/classification, problem solving, and explaining results. Each dimension is rated on a 5-point scale (1 = observed at a very limited level, 2 = limited, 3 = moderate, 4 = good, 5 = very good). Every child was rated on all seven dimensions at the end of each activity, and the overall activity score was the mean of the seven ratings. Child observation scores reported in the Findings are therefore on a 1–5 metric.

2.4.2. Activity Evaluation Observation Form

This form was used to assess the classroom applicability of each activity. Its criteria included appropriateness of materials, child participation, clarity of instructions, adequacy of time, story–STEM integration, classroom organization, safety, teacher/researcher guidance, and the closing–sharing phase, each rated on a 5-point scale (1 = inadequate, 2 = needs improvement, 3 = moderately adequate, 4 = good, 5 = very good). Scores informed descriptive judgments of applicability, and the accompanying observation notes were interpreted during thematic analysis under the headings of material appropriateness, child participation, clarity of instructions, adequacy of time, and story–STEM integration.

2.4.3. Semi-Structured Teacher Interview Form

Teacher interviews explored how the activities were reflected in children's scientific process skills, curiosity, problem-solving strategies, age-related differences, and classroom applicability. The form contained seven open-ended questions on children's observation, prediction, inference,

questioning, material exploration, and solution development; age-related differences; implementation facilitators and challenges; and suggestions for curriculum integration. Interviews were conducted individually, audio-recorded with consent, and transcribed verbatim.

2.4.4. Field notes

During each activity, the researcher kept structured field notes on children's use of materials, peer interactions, questioning, responses to problem situations, teacher interventions, and the flow of the session. Field notes were used to explain patterns found in the observation form and to cross-check teacher interview data.

2.4.5. Children's Products

Towers, bridges, pathway arrangements, door mechanisms, color-mixing trials, light–shadow set-ups, optical observation drawings, and other design products were examined through qualitative document analysis (Bowen, 2009). Products were evaluated for fitness for purpose, functionality, reflection of scientific understanding, redesign, and quality of problem solving. They were treated not merely as physical artifacts but as evidence that made children's inquiry, material use, solution generation, and design improvement visible.

2.5. Pilot Implementation and Instrument Development

The child observation form, the activity evaluation form, the semi-structured teacher interview form, and the children's product analysis matrix were piloted before the main implementation. The pilot involved six children who were not part of the main study, two from each age group, so that all three age bands were represented. Preliminary interviews were also held with two preschool teachers to check the clarity, sequencing, and applicability of the interview form.

The pilot showed that instructions for the 3–4 group in particular needed to be shorter, more concrete, and given one step at a time. Visual supports were therefore increased, material sets were reorganized by age group, and repetitive wording in the interview form was simplified. Pilot data were not included in the main analyses.

For content validity, two specialists in preschool education and one specialist in measurement and evaluation reviewed the instruments. They judged the items for age appropriateness, observability, clarity, and relevance to the research questions. Following their feedback, some items were simplified, behavioral indicators were made more explicit, and instructions were differentiated by age group.

2.6. Validity and Reliability

Validity and reliability were examined with procedures suited to each data source: expert review, pilot testing, internal consistency analysis, inter-observer reliability, inter-coder agreement, data triangulation, and researcher reflexivity.

For the STEM–Scientific Inquiry Observation Form, content validity was supported by the three experts described above, whose feedback led to simplified items, merged overlapping expressions, and clearer behavioral indicators. The form was piloted with six children outside the main sample. Internal consistency was high (Cronbach's $\alpha = .89$; McDonald's $\omega = .91$). A second expert independently scored 30% of the observation data, and inter-observer agreement was good (intraclass correlation coefficient [ICC] = .84, 95% confidence interval [CI] [.78, .89]).

The Activity Evaluation Observation Form was reviewed by the experts for content and face validity; some criteria were simplified and explanatory indicators clarified. Internal consistency was $\alpha = .87$ and $\omega = .88$.

The interview form was refined through expert review and the two preliminary teacher interviews, after which repetitive items were removed, the order of questions was revised, and probing questions were added. A second researcher independently coded 30% of the transcripts, yielding $\kappa = .84$, in the "almost perfect" range (Landis & Koch, 1977). For children's products, a

second coder rated 30% of the material, and agreement was κ = .81 (see Analysis of Children's Products).

To strengthen credibility, findings from all five sources were compared through triangulation and brought together in a joint display (see Convergence Across Data Sources). Member checking, expert review, thick description, an audit trail, and a reflective research journal were also used, in line with qualitative trustworthiness criteria (Lincoln & Guba, 1985). Taken together, these procedures provide acceptable evidence of validity and reliability within the scope of the study. Because the STEM–Scientific Inquiry Observation Form was developed for this specific context, however, further psychometric validation with larger samples is needed.

2.7. Data Collection Procedure

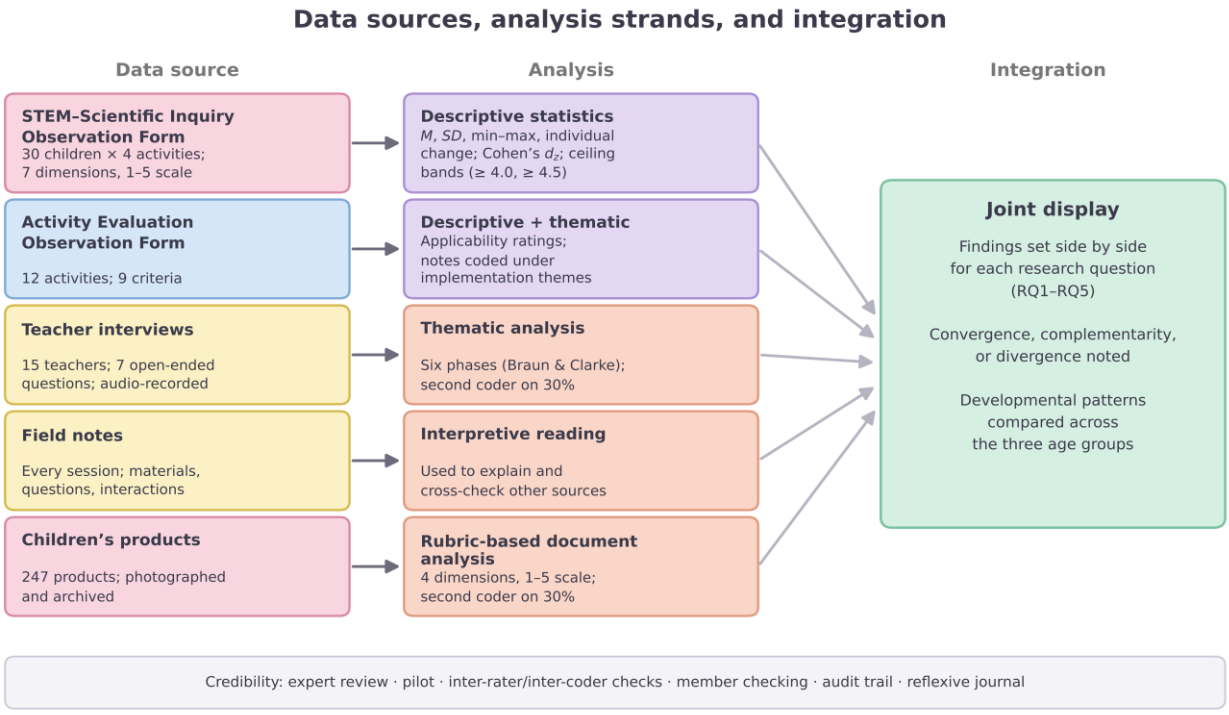
Data were collected during and after the four-week implementation. At the end of each activity, child observation forms and the activity evaluation form were completed and field notes were written. Products were photographed, classified, and archived for analysis. After the implementation ended, semi-structured interviews were held with the 15 teachers. Children were coded C1–C30 and teachers T1–T15 throughout.

2.8. Data Analysis

Figure 2 summarizes how each data source was analyzed and how the strands were brought together.

Figure 2

Data sources, analysis strands, and integration through a joint display



2.8.1. Descriptive statistics and effect size

Observation scores were analyzed by age group, implementation week, and skill dimension. Means, standard deviations, minimum–maximum values, and the standard deviation of child-level change scores were calculated for each age group. Given the small sample and the naturalistic, design-based character of the study, significance tests were not used as the primary analytic strategy.

Standardized effect sizes were computed to describe the magnitude of change. For comparisons between the first activity (A1) and the fourth activity (A4), Cohen's d_z was calculated from the

standard deviation of child-level difference scores. Ceiling tendency was examined descriptively to see whether children with high initial performance on the 5-point scale had limited room to improve; for this purpose, the proportion of children scoring 4.0 or above and 4.5 or above at the final activity was reported for each age group. This analysis was intended to indicate descriptively whether high scores constrained the observed change, not to test a ceiling effect formally.

2.8.2. Thematic analysis

Teacher interviews were analyzed following the six phases of thematic analysis described by Braun and Clarke (2006): familiarization with the data, generation of initial codes, searching for themes, reviewing themes, defining and naming themes, and reporting. Statements were first divided into meaning units; similar units were coded, and related codes were grouped into subthemes. The final themes were scientific process skills, curiosity and inquiry, problem-solving strategies, developmental differentiation, pedagogical and operational conditions, and curriculum and future practice. In line with Braun and Clarke's later account of reflexive thematic analysis (Braun & Clarke, 2019, 2021), themes were treated as patterns of meaning constructed by the researcher, and the researcher's background in early childhood STEM education was treated as an interpretive resource rather than a bias to be removed. The second-coder check described under Validity and Reliability was used to prompt discussion and refine code definitions; its agreement statistic is reported for transparency rather than as proof of coding accuracy.

2.8.3. Analysis of children's products

Products were analyzed with an analytic rubric developed for this study. In total, 247 products were examined, including towers, bridges, pathway arrangements, door mechanisms, color-mixing trials, light–shadow set-ups, optical observation drawings, and other design artifacts. Each product was first classified by age group, activity week, and product type and then rated on four dimensions: originality, functionality, scientific understanding, and problem solving/design revision (see Table 5).

Each dimension was rated on a 5-point scale, from 1 (very limited evidence) to 5 (highly visible and developmentally advanced evidence). Ratings did not rest on the final appearance of the product alone; each product was interpreted together with field notes, the child's verbal explanations, observed trial-and-error behavior, and evidence of revision during the activity. Products were thus treated as process-based evidence of inquiry, material use, design thinking, and problem solving.

To check reliability, 30% of the products were independently coded by a second coder with expertise in preschool education. Agreement was $\kappa = .81$, which falls in the "almost perfect" band of Landis and Koch (1977). Disagreements were resolved through discussion, and the agreed coding framework was applied to the remaining products. Because the same children produced several products across activities, product scores were used descriptively to compare developmental patterns and were not treated as independent observations in any inferential analysis.

2.9. Ethical Considerations

Ethical approval was obtained from an institutional ethics committee before data collection began. Written parental consent and children's age-appropriate verbal assent were obtained, and teachers gave written consent for participation and for audio-recording of interviews. All data were anonymized with the codes C1–C30 and T1–T15. Photographs and recordings were used only for research purposes, and no identifiable images appear in this article. Participants were informed of their right to withdraw at any time. The study complied with the Turkish Personal Data Protection Law No. 6698.

Table 5
Analytic rubric used to evaluate children's products

Dimension	*Operational Definition*	*Low-Level Evidence*	*Moderate-Level Evidence*	*High-Level Evidence*
Originality	The extent to which the product reflects the child's own idea, variation, or creative solution within the story-based STEM task	Product largely imitates the given example or uses materials without a clear idea	Product includes some individual choice or variation	Product reflects a distinctive, child-generated solution or creative adaptation
Functionality	The extent to which the product fulfills the intended task or problem requirement	Product does not work or is not connected to the task goal	Product partially fulfills the task but has limited stability or usability	Product works for the intended purpose and demonstrates task suitability
Scientific understanding	The extent to which the product reflects relevant STEM concepts such as balance, size, measurement, light, shadow, magnetism, or structure	Limited or no observable connection to the target concept	Partial evidence of concept use through material choice or arrangement	Clear evidence of concept use, often supported by explanation or purposeful design
Problem solving/design revision	The extent to which the child tests, modifies, improves, or justifies the product during the activity	No visible revision or persistence after failure	Some modification after trial or adult prompt	Purposeful testing, redesign, strengthening, or justification of the solution

3. Findings

This section reports the findings from the STEM–Scientific Inquiry Observation Form, teacher interviews, the Activity Evaluation Observation Form, field notes, and children's products in relation to the five research questions. Child scores are on a 5-point scale, and the overall observation score is the mean of the seven dimensions described in the Method section.

3.1. Change in Scientific Process Skills (RQ1)

The first research question (RQ1: What patterns in children's scientific process skills were observed across the activities?) was addressed mainly through the observation form. Over the four weeks, the overall observation score, the mean of the seven dimensions, increased in all three age groups (see Table 6).

Table 6
Activity-based means, standard deviations, and change indicators by age group

Age Group	*A1 M ± SD*	*A4 M ± SD*	*Mean Increase*	*95% CI*	*Cohen's d_z*
3–4 years	2.11 ± 0.45	2.98 ± 0.51	+0.87	[0.53, 1.21]	1.85
4–5 years	2.70 ± 0.38	3.80 ± 0.44	+1.10	[0.76, 1.44]	2.31
5–6 years	3.30 ± 0.32	4.26 ± 0.36	+0.96	[0.62, 1.30]	2.02

Note. A1 = first activity; A4 = fourth activity. Scores are on a 5-point rating scale; each age group included n = 10 children. CI = 95% confidence interval for the mean increase from A1 to A4. Cohen's d_z describes the magnitude of within-group change and should not be read as evidence of a causal effect or of retention.

All three groups showed large within-group changes, with the largest mean gain in the 4–5 group. Because the oldest group started highest, its final mean was also the highest even though its gain was smaller than that of the 4–5 group.

Teacher interviews pointed in the same direction. T1 noted that children initially examined materials mostly by surface features such as color and shape, but later began to use more functional criteria, asking, for example, "Is it wide enough for Alice to pass through?" This suggests that observation and comparison became more purposeful over time. Looking across the seven dimensions, the change was not confined to overall scores; children displayed a range of distinct indicators of scientific inquiry (see Table 7).

Table 7

Observed patterns of change across the dimensions of the STEM–Scientific Inquiry Observation Form

Dimension	*Observed Pattern of Change*	*Data Source*	*Evidence*
Observation	Children began to examine materials not only by surface features such as color, shape, and size but also by their functional properties.	Observation form, teacher interviews, field notes	"This one is big, but this one is strong" (C4). "This door is too small for Alice; we need a bigger one" (C13). "When I press this block, it stays straight, so it is better for the bridge" (C24).
Prediction	As the activities progressed, children gave more answers to "What will happen?" and increasingly linked their predictions to trials.	Observation form, field notes	"If we put it here, maybe it will fall" (C5). "If the road is longer, Alice can go farther" (C16). "If we add one more support, the bridge will not collapse" (C27).
Inference	Children began to relate outcomes to their earlier trials and to explain why a result occurred.	Field notes, children's products	"It fell because this side was little" (C8). "This one did not open because the paper was stuck" (C18).
Comparison	Children compared materials by size, durability, length, width, and function.	Observation form, children's products	"This one is big, this one is small" (C2). "This key is longer, but that one is thinner" (C14). "The cardboard bridge is weaker than the wooden one" (C23).
Measurement/ classification	Children began to sort objects by size, color, length, width, and intended use.	Observation form, field notes	"Big blocks here, small blocks there" (C7). "These are the red ones, these are the blue ones" (C12).
Problem solving	Children changed ineffective solutions, tried different materials, and redesigned their products.	Observation form, children's products, teacher interviews	"This does not work; let's try this one" (C9). "We need another block so it can stand" (C17). "I changed the middle part because it was not holding the weight" (C29).
Explaining results	Especially in the 5–6 group, children said more clearly what they had done and why they had chosen a particular solution.	Teacher interviews, field notes, children's products	"I made it like this because Alice needs a bigger chair" (C10). "I chose this material because it was stronger, and that is why the bridge stayed up" (C28).

Note. The table integrates the seven dimensions of the observation form with evidence from teacher interviews, field notes, and children's products. Child codes: C1–C10 = 3–4 years; C11–C20 = 4–5 years; C21–C30 = 5–6 years.

As Table 7 shows, the change in scientific process skills was visible not only in rising scores but in what children actually did. Across the activities, attention moved from surface features of materials, such as color, size, and shape, toward functional properties, including strength, width, stability, and suitability for the task. Predictions and inferences also became more purposeful as children tied their ideas to the outcomes of trials and earlier attempts.

Comparison, measurement/classification, and problem solving surfaced in children's efforts to sort materials, test alternatives, modify solutions that did not work, and redesign their products. These behaviors appeared in every age group, although the 5–6-year-olds gave the most explicit explanations for their design choices and outcomes. Read together, the observation forms, field notes, teacher interviews, and products indicate that changes occurred across several dimensions of scientific inquiry, in developmentally differentiated ways (see Table 8).

Table 8
Weekly micro-patterns by STEM–Scientific Inquiry dimension

Dimension	*A1*	*A2*	*A3*	*A4*
Observation	Focusing on color, shape, and size of the material	Distinguishing material properties	Comparing functional properties	Linking observation to the solution
Prediction	Simple, brief predictions	Predicting before testing	Testing predictions with materials	Comparing predictions with outcomes
Inference	Noticing the outcome	Relating the outcome to previous trials	Establishing cause–effect connections	Justifying one's solution
Comparison	Distinguishing big/small and long/short	Classifying materials by durability and suitability	Comparing solution strategies	Selecting the most functional solution
Measurement/ classification	Sorting objects by basic properties	Grouping by size and intended use	Distinguishing suitable from unsuitable materials	Turning classification into design decisions
Problem solving	Trial and error	Trying alternative materials	Redesigning the product	Testing and improving the design
Explaining results	Brief statements about outcomes	Stating what happened	Attempting to explain why it happened	Justifying one's solution

Note. A1–A4 = activities in weeks 1–4.

Table 8 indicates that the dimensions of inquiry became visible through different behaviors depending on the structure of each activity, rather than through a uniform weekly rise. In the first week, children were mostly getting to know the materials, making basic observations, and offering simple predictions. By the third and fourth weeks, testing predictions, revising solutions, and explaining results were more prominent. The findings were therefore read not only as changes in mean scores but as shifting patterns of observable skill across the implementation.

3.2. Transformation in Problem-Solving Strategies (RQ2)

For RQ2 (What changes in children's problem-solving strategies became visible during the activities?), scores on the problem-solving dimension of the observation form rose in all three groups, most in the 4–5 group. These increases are consistent with children approaching the story-based problems more purposefully over time. To see how this shift differed by age, the score changes were read alongside teacher interviews, field notes, and products (Table 9).

Table 9
Transformation in children's problem-solving strategies across age groups

Age Group	*Increase in Problem-Solving Score*	*Early-Stage Pattern*	*Later-Stage Pattern*	*Supporting Evidence*
3–4 years	+0.90	Mainly trial and error and direct manipulation of materials. Solutions were often immediate and based on visible features of the materials.	Children began to try alternative materials, repeat attempts, and make simple modifications when a solution did not work.	Products showed repeated re-stacking and rearranging of blocks. Field notes indicated growing persistence after failure; for example, "This one fell; I will try the other one" (C6).
4–5 years	+1.20	Children began to recognize when a solution was ineffective and tried different approaches, though not yet systematically.	Children more often generated alternative solution paths, compared materials, and revised their products to improve functionality.	Teacher interviews and field notes showed more purposeful testing of alternatives, e.g., "We need another block so it can stand" (C17). Products reflected multiple solution attempts, especially in pathways and doors.
5–6 years	+1.00	Children started with more planned actions than younger groups, but first solutions were not always stable or fully functional.	Children increasingly tested, evaluated, and redesigned their products, adding supports, strengthening structures, and justifying revisions verbally.	Teacher interviews, field notes, and products showed a design-oriented approach. One child explained, "I changed the middle part because it was not holding the weight" (C29). Products showed revision, reinforcement, and justification.

Note. Problem-solving scores are from the problem-solving dimension of the STEM–Scientific Inquiry Observation Form (A1 to A4). Child codes: C1–C10 = 3–4 years; C11–C20 = 4–5 years; C21–C30 = 5–6 years.

The patterns in Table 9 suggest that the change in problem-solving strategies was developmental in character. For the youngest children, problem solving consisted mostly of trial and error, repeated attempts, and direct manipulation. The 4–5-year-olds began testing different solution paths and revising their products to make them work better. Among the 5–6-year-olds, problem solving took a more design-oriented form: children tested their products, located weak parts, added supports, and explained why they had made changes. T5 described how children who had previously asked an adult for help as soon as their towers collapsed later began to try different pieces and make their structures stronger. Field notes and products showed the same shift, particularly in bridge, tower, and door-mechanism activities. Overall, the evidence describes a progression from immediate trial and error toward more intentional, reflective, and design-based problem solving.

3.3. Developmental Differences and Ceiling Tendency (RQ3)

For RQ3 (How did children in the three age groups respond to the same narrative framework?), Table 10 shows that the age groups differed both in where they started and in where they finished.

The youngest group began in the lowest range and remained below the high-performance band at the end, which suggests that 3–4-year-olds showed gains but continued to need concrete, guided

Table 10
Score ranges, variability of change, and ceiling tendency by age group

Age Group	*A1 Min–Max*	*A4 Min–Max*	*SD of Change*	*A4 ≥ 4.0 n (%)*	*A4 ≥ 4.5 n (%)*
3–4 years	1.5–2.9	2.2–3.9	0.47	0 (0%)	0 (0%)
4–5 years	2.2–3.2	3.2–4.4	0.48	3 (30%)	0 (0%)
5–6 years	2.8–3.7	3.8–4.7	0.48	8 (80%)	2 (20%)

Note. SD of change = standard deviation of child-level A1-to-A4 difference scores (mean change divided by Cohen's d_z; see Table 6). A4 ≥ 4.0 marks the high-performance band; A4 ≥ 4.5 indicates proximity to a ceiling tendency.

and carefully scaffolded experiences. Variability in individual change was similar across the three groups, so the groups differed less in how much children varied than in where their gains were located. The 4–5 group showed the largest mean gain and included three children who reached the high-performance band, a profile consistent with a transitional period in which more systematic inquiry and problem solving begin to appear. The 5–6 group started higher and finished in the highest range: 80% scored 4.0 or above at A4, which points to more purposeful observation, comparison, explanation, and design revision. Only 20% reached the 4.5 threshold, however, so the pattern reflects high performance rather than a strong ceiling effect. In sum, the same story-based framework was associated with developmentally differentiated response patterns, with younger children growing mainly through material exploration and older children showing more advanced, design-oriented inquiry.

3.4. Curiosity and Exploration Behaviors (RQ4)

RQ4 asked through which observable behaviors curiosity and exploration appeared. Curiosity and exploration were not scored as a separate subdimension of the observation form. They were evaluated instead through observable behaviors recorded in teacher interviews, field notes, the Activity Evaluation Observation Form, and children's products, with material exploration, questioning, prediction, and attempts to test different solution paths considered together. Table 11 summarizes these patterns by age group.

Table 11
Age-related patterns of curiosity and exploration behaviors

Age Group	*Main Focus of Curiosity*	*Observable Behaviors*	*Data Source*	*Evidence*
3–4 years	The story problem and concrete materials	Touching and examining materials, repeating simple trials, asking short questions, reacting to visible changes	Teacher interviews, field notes, Activity Evaluation Observation Form	Curiosity was often triggered by immediate story problems and the physical properties of materials, e.g., "Will Alice fit here?" (C3); "Can I try this one too?" (C7).
4–5 years	How materials could be used and whether different solutions might work	Trying alternative materials, comparing options, asking "what if" questions, modifying attempts	Teacher interviews, field notes, children's products	Children more often explored alternative uses of materials and solution paths, e.g., "What happens if we make the road longer?" (C15); "Let's try another piece; maybe it will open" (C18).
5–6 years	Causal explanations, design improvement, and the logic of outcomes	Testing products, asking why a solution worked or failed, revising designs, explaining results	Teacher interviews, field notes, children's products	Curiosity took a more analytical form, e.g., "Why did the bridge fall even though it was long?" (C24); "If we put support under it, it will hold better" (C29).

Note. Curiosity and exploration were not scored quantitatively; they were analyzed through teacher interviews, field notes, the Activity Evaluation Observation Form, and children's products. Child codes: C1–C10 = 3–4 years; C11–C20 = 4–5 years; C21–C30 = 5–6 years.

Curiosity and exploration appeared in every age group but took developmentally different forms. Among 3–4-year-olds, curiosity was expressed mainly through direct engagement with materials and brief, story-related questions. In the 4–5 group, it became more closely tied to exploring alternative uses of materials and testing different solution paths. By 5–6 years, curiosity had a more analytical, design-oriented quality: children asked why a product worked or failed, considered modifications, and tried to explain outcomes.

Teacher interviews supported this reading. T1 observed that children examined materials for longer in scenes where Alice faced a problem, suggesting that the narrative helped sustain attention and exploration. Field notes recorded questions such as "Will this work?", "Are there more materials?", and "What happens if we do it like this?", indicating that curiosity extended beyond interest in the story itself to sustained exploration of materials, strategies, and results. Curiosity thus appeared not merely as initial engagement with the narrative but as an evolving inquiry behavior that became more purposeful, and more differentiated by age, over the activities.

3.5. Developmental Analysis of Children's Products

Children's products served as a complementary source of evidence on scientific process skills, problem-solving strategies, and developmental differences. The analysis covered 247 products rated with the rubric in Table 5 (see Table 12).

Table 12
Developmental comparison of children's products by evaluation dimension

Dimension	*3–4 Years*	*4–5 Years*	*5–6 Years*	*Developmental Pattern*
Originality	2.60	3.30	3.90	Originality increased with age.
Functionality	2.40	3.20	3.80	Products were more functional in the older groups.
Scientific understanding	2.10	3.00	3.70	The most visible age-related difference emerged in this dimension.
Problem solving/design revision	2.30	3.20	3.90	Design revision was particularly evident in the 5–6 group.
Overall mean	2.35	3.18	3.83	Product quality increased gradually across age groups.

Note. Products were rated on a 5-point analytic rubric; values are descriptive mean ratings by age group. Inter-coder agreement: $\kappa = .81$.

The product analysis offered concrete evidence of differentiation across ages. Products from the 3–4 group were characterized mainly by direct manipulation, simple construction attempts, and trial and error. Although these products were often less stable or elaborate, they documented early exploration and growing engagement with the task.

In the 4–5 group, products showed a clearer move toward purposeful construction. Children compared materials, tried alternative arrangements, and improved the functionality of what they built, particularly in the pathway and door activities, where they worked to fit their products to the story problem.

Products from the 5–6 group showed the strongest evidence of scientific understanding and design revision. These children more often tested their products, identified weak or non-functional parts, added supports, reorganized materials, and justified their decisions aloud. The age-related rise in product scores was therefore read not simply as a difference in final quality but as evidence of increasingly sophisticated inquiry, problem solving, and design reasoning.

3.6. Classroom Applicability (RQ5)

RQ5 concerned teachers' evaluation of classroom applicability. Most teachers considered that the story-based STEM model could be integrated into the preschool curriculum. Twelve teachers

described the activities as compatible with daily plans and interdisciplinary integration, and 14 emphasized that ready-to-use activity plans, material lists, and sample instructions would be needed for sustainable implementation (see Table 13).

Table 13
Teacher views on the classroom applicability of the activities

Theme	*Codes*	*f*	*Teachers*	*Representative Statement*
Curriculum integration	Alignment with daily plans, interdisciplinary integration, story–science–mathematics connection	12	T1–T4, T6–T9, T11–T14	"Integrating science, mathematics, and design tasks through the story seems adaptable to the daily plan." (T8)
Need for ready-to-use activity plans	Age-specific planning, material lists, sample instructions, learning outcome alignment	14	T1–T5, T7–T15	"These activities are applicable; however, the process would be more sustainable if teachers had ready-to-use plans and material lists." (T12)
Visual and material support	Visual cards, modeling, durable materials, age-appropriate tools	11	T1–T3, T5–T8, T10, T12, T14, T15	"Especially for the younger age group, visual support and preparing the materials in advance were very important." (T5)
Time and classroom organization	Small-group work, station arrangement, flexible time, structured closure	10	T1, T3, T4, T6, T8–T11, T13, T14	"When children worked in small groups, they were able to experiment more comfortably." (T10)
Teacher guidance	Open-ended questioning, supporting without directing, researcher/teacher coordination	14	T1–T6, T8–T15	"Guiding children through questions without directly providing the solution improves the quality of the implementation." (T3)

Note. *f* = number of teachers (out of 15) whose interviews contained the theme.

Teachers thus regarded the activities as applicable in classrooms, but only under certain pedagogical and operational conditions. Age-appropriate instructions, visual supports, materials prepared in advance, small-group organization, a structured closing phase, and teacher guidance were all seen as raising the quality of implementation. Story-based STEM activities, in this view, need to be planned not only for content but also for classroom organization and teacher support.

3.7. Convergence across Data Sources

To integrate the findings, a joint display was constructed (Guetterman et al., 2015) in which results from the observation form, teacher interviews, the activity evaluation form, field notes, and children's products were set side by side for each research question (see Table 14). The display shows largely convergent, mutually supportive, and complementary patterns across sources.

Across the five research questions, quantitative observation data, teacher interviews, field notes, and products told a coherent story. Children's inquiry, problem solving, curiosity, and exploration changed during the activities in ways that differed by age. The quantitative data documented rising inquiry and problem-solving scores, and the qualitative evidence showed how these changes played out in practice through children's explanations, material exploration, prediction testing, collaborative attempts, and product revisions.

Table 14
Joint display of findings across data sources by research question

Research Question	*Observation Form*	*Teacher Interviews*	*Field Notes and Products*	*Convergence Pattern*
RQ1: Patterns in scientific process skills	Increases from A1 to A4 in all age groups	Teachers reported more observation, prediction, and explanation	Children tested predictions, explained products, and shared results	Changes in scientific process skills supported by multiple sources
RQ2: Changes in problem-solving strategies	Increases in the problem-solving dimension in all age groups	Teachers described retrying, generating alternatives, and revising designs	Products showed added supports, changed mechanisms, and redesigned structures	Problem solving extended from trial and error toward design revision
RQ3: Responses of the three age groups	Largest increase in the 4–5 group; highest final scores in the 5–6 group	Concrete material use more prominent among younger children; explanation and justification among older children	Complexity and functionality of products increased with age	A consistent developmental progression across age groups
RQ4: Curiosity and exploration	Not scored separately; associated with inquiry-related behaviors during activities	Children asked more questions and examined materials for longer	Questions such as "What happens if we do it like this?" recorded	Story context identified by teachers and field notes as linked to curiosity and exploration
RQ5: Classroom applicability	Inquiry and problem-solving processes were observable and manageable in the classroom	Activities applicable, given planning, materials, and guidance	Small-group work, preparation, and structured closure eased implementation	Applicable in classrooms when age-appropriate adaptations are made

The changes, then, were not limited to numerical gains; they were also visible in classroom behavior. Children increasingly compared materials, tested predictions, revised designs, explained outcomes, and engaged more purposefully with the problems embedded in the story. At the same time, teacher views and field notes made clear that successful implementation depended on pedagogical scaffolding, age-appropriate instructions, material preparation, small-group organization, and guided questioning. These integrated findings form the basis for the discussion that follows of the pedagogical value, developmental implications, and practical conditions of story-based STEM in preschool education.

4. Discussion

This study examined how story-based STEM activities built around *Alice's Adventures in Wonderland* were reflected in preschool children's scientific process skills, problem-solving strategies, curiosity and exploration, and age-related responses. Positive changes in scientific process and problem-solving behaviors were observed in every age group, but they followed different patterns by age. The story context appeared to function as a pedagogical element associated with curiosity and exploration, and teachers regarded the activities as applicable in

their classrooms. The sections below discuss these findings in light of the literature, propose design principles, and then consider alternative explanations.

4.1. Change in Scientific Process Skills

With respect to RQ1 (patterns in scientific process skills), scores rose in all age groups over the four weeks, with descriptively large within-group effect sizes (d_z = 1.85–2.31). This pattern fits work showing that scientific thinking in early childhood can develop through interaction with concrete materials, observation, prediction, and explanation of outcomes. Gelman and Brenneman (2004) argue that, given suitable environments, young children can use observation, comparison, and causal inference. Eshach and Fried (2005) similarly maintain that early science should be supported through exploration, experience, and the production of explanations rather than through direct transmission of concepts. The more purposeful examination of materials, testing of predictions, and explanation of outcomes observed here, when children met problems embedded in the story, is consistent with this literature. In psychological terms, these behaviors can be read as early instances of the hypothesis–evidence coordination that Zimmerman (2007) places at the heart of scientific reasoning.

The dimension-level analyses show that the change was not one-dimensional. Observation shifted from surface description toward functional examination and comparison, and predictions developed from simple statements into more justified, testable claims. McClure et al. (2017) argue that when preschool STEM outcomes are framed as inquiry-based process skills rather than as domain-specific knowledge alone, they can be carried more meaningfully into children's learning experiences. The dimensional patterns found here suggest that story-based STEM activities gave children opportunities to make such process skills visible.

4.2. Transformation in Problem-Solving Strategies

With respect to RQ2 (problem-solving strategies), scores rose in all three age groups. When children met problems embedded in the story, they modified ineffective solutions, tried alternative materials, and redesigned their products, so problem solving became a process rather than a single step. This is in line with studies reporting that design thinking and STEM-based activities in preschool can support early problem-solving, creative-thinking, and engineering-design skills (Akcay Malcok & Ceylan, 2022; Yalçın & Erden, 2021).

The behavioral changes also suggest that the story acted as more than an attention-getter; it seems to have helped children make sense of the problem and build solutions within the flow of the narrative. Marcus et al. (2021) found that narrative reflection can strengthen children's STEM talk and later recall. The design explanations and justifications seen especially among 5–6-year-olds suggest that a story context may help children articulate their problem-solving, although transfer to new problems was not assessed here. Yang et al. (2023) similarly showed that stories place children's task planning and solution generation within a meaningful context. The persistence that teachers described after failed attempts also recalls an earlier preschool STEM program in which perseverance was among the qualities teachers most associated with children's progress (Şahin, 2026a). From a scaffolding perspective (Wood et al., 1976), the narrative structure may have carried part of the task structure, such as the goal and the constraint, leaving children free to concentrate on testing and revising.

4.3. Age-Related Developmental Differentiation and Ceiling Tendency

With respect to RQ3 (responses of the three age groups), all three age groups responded positively, but the form and size of their responses differed. Gains among 3–4-year-olds were concentrated in concrete material use, trial and error, and basic classification, while the 4–5 group showed the largest mean increase. In the 5–6 group, children clustered in the high-performance band (80% at or above 4.0 at A4), and some approached a descriptive ceiling (20% at or above 4.5). Groups starting from higher levels may simply have less room to improve, and scores may cluster near the upper thresholds of a single observation instrument.

This pattern is consistent with research showing that STEM and scientific thinking skills vary with age in symbolic representation, language, and cognitive flexibility (Vidal Carulla et al., 2021). Uyanık et al. (2026), who reported improvements in preschool children's scientific process skills following teacher training in story-based 5E science instruction, also highlighted the value of age-appropriate story content and suitable assessment tools. The ceiling tendency observed among the oldest children in the present study suggests that story-based STEM research should attend not only to changes in mean scores but also to the proportion of children reaching high-performance bands and the distribution and range of individual scores.

4.4. Making Curiosity and Exploration Visible

With respect to RQ4 (curiosity and exploration), children examined materials, asked questions, tried alternative solution paths, and approached problems by identifying with the story characters. These behaviors suggest that the story context made curiosity and exploration visible within the classroom. The finding fits the view of curiosity and wonder as powerful sources of scientific learning in early childhood (Bjerknes et al., 2024) and Loewenstein's (1994) account of curiosity as arising when a person becomes aware of a specific gap in knowledge; a character who cannot fit through a door creates exactly such a gap. That curiosity indicators converged across products, teacher interviews, and field notes indicates that story-based STEM activities can make these behaviors observable through multiple sources.

4.5. The Story as a Pedagogical Context

The story appears to have functioned not merely as an engaging introduction but as a pedagogical context that helped children make sense of problems, experiment with materials, and explain their solutions. Cremin et al. (2017) describe narrative and storytelling in early childhood as a fundamental resource not only for language and literacy but also for classroom culture, thinking, and engagement, and Cremin et al. (2015) note that early scientific learning is supported by play, creativity, dialogue, and inquiry. Consistent with the research on learning from fiction reviewed earlier (Hopkins & Weisberg, 2017; Walker et al., 2015), the fantastical events seem to have worked as prompts for questions rather than as sources of facts. Turning events from *Alice's Adventures in Wonderland* (growing and shrinking, the lost key, passing through the door, reflection, shadow, and balance) into problem situations for STEM tasks suggests that a story's narrative structure can offer children both a conceptual context and sustained motivation. This interpretation is compatible with Bruner's (1986) view of narrative as a mode of thought that organizes events around intentions and obstacles.

This interpretation also fits literature suggesting that narrative-based approaches can support conceptual contextualization, task continuity, and recall in preschool STEM (Altun, 2019; Miller-Goldwater et al., 2024), and that narrative, digital storytelling, and gamified practice can benefit preschool children's learning interest, scientific concept development, and problem solving (Eliza et al., 2025; Pahl et al., 2025; Shemy, 2020; Vaknin-Nusbaum & Nevo, 2017; Zhou & Yang, 2025). Many of these studies, however, do not systematically report performance distributions, the share of children reaching high-performance bands, or the extent to which high starting levels limit later improvement. By reporting individual score ranges and indicators of ceiling tendency, the present study shows why story-based preschool STEM research should consider how different age groups approach the upper limits of performance, not only mean increases.

4.6. Classroom Applicability and Teacher Perspectives

With respect to RQ5 (classroom applicability), teachers considered the activities applicable but stressed that sustainable implementation requires ready-to-use plans, age-appropriate materials, visual supports, small-group organization, structured closure, and teacher guidance. This echoes studies in Türkiye that highlight the importance of teacher support, material resources, and professional development in integrating STEM into preschool classrooms (Uğraş & Genç, 2018; Uğraş et al., 2025). That teachers found the approach acceptable while also naming needs for

planning, material preparation, and professional development suggests that story-based STEM should be structured not only at the level of activity design but also at the level of classroom arrangements. The request for ready-to-use plans also mirrors the experience of child development students, who reported difficulty planning STEM activities until they had gained the relevant content knowledge (Şahin, 2021), while the emphasis on guiding children without directing them resembles the emotionally responsive pedagogy described in an integrated STEM and social–emotional learning program (Şahin, 2026a). Similar calls for structured, practice-oriented support appear in recent work with Turkish preschool teachers on social and emotional learning (Özdemir Beceren et al., 2025), and the value of guiding without directing echoes participatory pedagogy, in which teachers learn to leave real choices to children (Akyol, 2025). Resources that provide teachers with ready-to-use content, visual supports, and implementation guidance may be important for sustaining such practice (Mohana et al., 2022).

4.7. Design Principles for Age-Sensitive Story-Based STEM

Design-based research is expected to yield not only findings but also transferable design knowledge (Anderson & Shattuck, 2012; Barab & Squire, 2004). Drawing together the weekly revisions, the age-group patterns, and teachers' views, Table 15 states five provisional design principles. Each is tied to the evidence from which it was derived, and each should be treated as a conjecture to be tested in other settings rather than as an established rule.

Table 15
Provisional design principles derived from the study

Principle	*Description*	*Main Supporting Evidence*
1. Fantasy frames the question; materials answer it	Use the story to create a puzzling situation, then restate it as a question children can settle by testing real materials.	Intervention design (see Story-Based STEM Activities); children's story-linked questions and functional comparisons (Tables 7 and 11)
2. One episode, one problem	Reduce each session to the single episode that sets up the problem, so that the narrative does not compete with the task.	Pilot revisions (see Pilot Implementation and Instrument Development); weekly review of instructions and time (see Research Design)
3. Differentiate by age, not only by difficulty	Vary session length, instruction format, materials, and inquiry emphasis across age groups (Table 4).	Age-group score and product patterns (Tables 10 and 12); teacher views (Table 13)
4. Make revision and explanation part of the task	Build time for testing, redesign, and a closing phase in which children explain what happened.	Shift from trial and error toward design revision and justification (Tables 8 and 9)
5. Support teachers with ready-to-use resources	Provide plans, material lists, visual cards, and question prompts that guide without directing.	Teacher themes on plans, visual support, and guidance (Table 13)

4.8. Alternative Explanations for the Observed Changes

Several explanations other than the story-based activities themselves could account for part or all of the observed changes, and they should be weighed alongside the interpretations above. First, children's growing familiarity with the materials, the session routine, and the observer could raise scores over four weeks regardless of narrative content; practice and novelty effects of this kind cannot be separated from intervention effects without a comparison group. Second, the researcher designed and implemented the activities and also rated the children. Although 30% of

observations were double-scored with good agreement (ICC = .84), neither rater was blind to the implementation week, so expectancy may have contributed to rising ratings.

Third, A1 and A4 were different activities with different content and demands (Table 3). A change from A1 to A4 therefore mixes children's development with differences in task difficulty, and between-group comparisons mix age with differences in activity set and session length. The larger increase in the 4–5 group, for instance, might reflect a genuinely responsive developmental period, but it might equally reflect the particular activities assigned to that group or the scale's greater sensitivity in its middle range, compared with the compressed upper range reached by the oldest children. Fourth, because the activities combined a story with hands-on materials, guided questioning, and small-group work, the study cannot isolate the contribution of the narrative itself; a non-narrative version of the same STEM tasks would be needed for that. These alternatives do not invalidate the descriptive patterns, but they limit how far the patterns can be attributed to the story-based design.

5. Limitations

Five limitations are highlighted here, considered separately for the quantitative and qualitative strands because the two rest on different quality criteria.

In the quantitative strand, the first limitation is the small sample of 30 children, 10 per age group, so the observation scores describe these children rather than estimate population values; the subgroups were also too small to examine sex differences. Second, the lack of a control group, the four-week duration, and raters who were not blind to the implementation week mean that score changes cannot be attributed to the activities and that retention was not assessed (see Alternative Explanations for the Observed Changes). Third, the observation form was developed for this study and still requires psychometric validation with larger samples.

The qualitative strand was not designed to generalize or to control for alternative explanations but to build a situated, interpretive account. From the perspective of reflexive thematic analysis, the absence of a control group is not a weakness of the qualitative findings, and the researcher's combined roles as designer, implementer, and observer were an intentional source of insight into how the activities unfolded (Braun & Clarke, 2019, 2021). The fourth limitation is the other side of the same coin: these roles may have shaped what teachers chose to say about activities designed by their interviewer, so favorable views may partly reflect social desirability. Relatedly, inter-coder agreement statistics, reported here for transparency, sit more comfortably with codebook approaches than with reflexive thematic analysis. Fifth, the study took place in one summer-school setting; rather than claiming generalizability, the article provides contextual description so that readers can judge how far the interpretations transfer to their own settings.

6. Conclusion and Recommendations

This design-based study examined how STEM activities built around *Alice's Adventures in Wonderland* were reflected in preschool children's scientific inquiry, problem solving, curiosity–exploration, and age-related responses. Positive changes in scientific process skills and problem solving were observed in every age group, along different developmental paths. The largest mean increase in the 4–5 group may indicate a particularly responsive transitional period, although differences in activity content and scale sensitivity offer competing explanations. The clustering of 5–6-year-olds in the high-performance band, together with a descriptive ceiling tendency, points to the need for assessment tools that discriminate better at the upper end.

Across the study, children came to examine materials more purposefully, test their predictions, generate alternatives, revise their products, and explain outcomes, and the quantitative increases converged with this qualitative evidence. Such convergence is consistent with the view that story-based STEM activities offer an integrated context for preschool children's inquiry, problem solving, curiosity, and exploration, while falling short of evidence that the activities caused these changes.

The study adds to the early childhood STEM literature in four ways. It offers an applied model for turning a classic fantasy text into STEM activities, mapped onto national program outcomes, and shows how fantasy can serve as a problem frame without being treated as a source of facts. It documents that the same narrative framework met with different developmental responses at 3–4, 4–5, and 5–6 years. It illustrates how several complementary data sources can be combined to evaluate story-based practice. Finally, it condenses these results into five provisional design principles that other researchers and teachers can test.

Story-based STEM appears feasible as a pedagogical approach for supporting scientific inquiry, curiosity, and problem solving in preschool, and it warrants controlled testing. For the age groups within this study, implementation depended on activities differentiated by developmental level, materials and instructions prepared in advance, carefully planned time and grouping, and practical guidance for teachers.

6.1. Recommendations for Practice and Policy

- Given teachers' call for ready-to-use plans and material lists, implementation guides aligned with preschool learning outcomes, differentiated by age group, and supported with visual materials could be developed.
- For this age range, STEM activities can be grounded mainly in story, play, concrete materials, first-hand experience, and problem solving rather than in complex digital tools.
- Teacher education could go beyond introducing STEM concepts to include practice-oriented skills such as turning stories into problem situations, asking open-ended questions, selecting suitable materials, and preparing age-appropriate instructions.
- Large, tactile, easily manipulated materials can be prioritized for 3–4-year-olds; symbolic play and multipurpose materials for 4–5-year-olds; and testing designs, explaining outcomes, and justifying solutions for 5–6-year-olds.
- Elements of works such as *Alice's Adventures in Wonderland*, including size, balance, reflection, time, doors, keys, and transformation, can be translated into STEM concepts such as measurement, cause–effect reasoning, light and shadow, and engineering design.
- Short sharing and evaluation phases at the end of each activity give children the chance to test, explain, and revise their products.

6.2. Recommendations for Future Research

- Because the present findings describe within-group change, future studies could test the approach with comparison-group designs, ideally including a non-narrative version of the same STEM tasks, blinded raters, and follow-up measurement.
- Comparative research could examine how the model functions across school types, age groups, and cultural contexts, holding activity content constant across age groups.
- Preschool-appropriate measures of scientific process skills, curiosity, and problem solving in story-based STEM should be validated with larger samples, with attention to discrimination at the upper end of the scale.
- Later design cycles could adapt the activities and instruments for inclusive classrooms so that children receiving special education services can take part with individualized support.

Acknowledgements: The author thanks the children, families, and preschool teachers who participated in the study, and the school administration for facilitating the implementation process.

Data availability: The data are available from the corresponding author upon reasonable request.

Declaration of generative AI and AI-assisted technologies: A generative artificial intelligence tool (Claude Opus 5.5, Anthropic) was used under the author's supervision for language editing, for restructuring the manuscript during revision, and for drafting the plotting code used to draw Figures 1 and 2 from content specified by the author. No artificial intelligence tool was used in

designing the study, collecting or analyzing the data, or interpreting the findings. The author reviewed all output and takes full responsibility for the accuracy and integrity of the manuscript.

Declaration of interest: The author declares no conflicts of interest.

Ethical approval and consent to participate: Ethical approval for this study was obtained from the Social Sciences and Humanities Ethics Committee of Çanakkale Onsekiz Mart University (Application No. 2025-YÖNP-0648; Decision No. 30/56, dated 18 July 2025) before data collection began. Written informed consent was obtained from the parents of all participating children, and age-appropriate verbal assent was obtained from the children themselves. The participating teachers provided written informed consent for participation and for the audio recording of the interviews. All participants were informed of their right to withdraw at any stage. Consent for the publication of the anonymized data was also obtained. All participants were anonymised using the codes C1–C30 (children) and T1–T15 (teachers); no identifiable visual material is included in the manuscript.

Funding: This research received no external funding.

Previous presentation: This study, entitled "A Study on Fostering Scientific Inquiry and Curiosity Through Story-Based STEM Activities in Preschool Education," was presented as an abstract paper at the 24th International Symposium on Primary School Teacher Education, held in İstanbul, Türkiye, on 14–17 May 2026. The present manuscript is a substantially expanded and revised version of that abstract and has not been published elsewhere.

References

Akcay Malcok, B., & Ceylan, R. (2022). The effects of STEM activities on the problem-solving skills of 6-year-old preschool children. *European Early Childhood Education Research Journal*, *30*(3), 423–436. https://doi.org/10.1080/1350293X.2021.1965639

Akyol, T. (2025). The effect of participatory pedagogy and children's participation training on early childhood teachers' views on children's participation. *Journal of Pedagogical Research*, *9*(4), 102–118. https://doi.org/10.33902/JPR.202533781

Altun, D. (2019). From story to science: The contribution of reading fiction and hybrid stories to conceptual change with young children. *Children & Society*, *33*(5), 453–470. https://doi.org/10.1111/chso.12321

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25. https://doi.org/10.3102/0013189X11428813

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Bers, M. U. (2021). *Coding as a playground: Programming and computational thinking in the early childhood classroom* (2nd ed.). Routledge. https://doi.org/10.4324/9781003022602

Bjerknes, A.-L., Wilhelmsen, T., & Foyn-Bruun, E. (2024). A systematic review of curiosity and wonder in natural science and early childhood education research. *Journal of Research in Childhood Education*, *38*(1), 50–65. https://doi.org/10.1080/02568543.2023.2192249

Blanquet, E., & Picholle, É. (2012). Inquiry-based analysis of early years children books: Developing skills for later science education. In C. Bruguière, A. Tiberghien, & P. Clément (Eds.), *E-book proceedings of the ESERA 2011 conference: Science learning and citizenship* (Part 14, pp. 6–11). European Science Education Research Association. https://www.esera.org/esera-2011/

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

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

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Bruner, J. (1986). *Actual minds, possible worlds*. Harvard University Press. https://doi.org/10.4159/9780674029019

Cheng, M.-M., & Chuang, H.-H. (2019). Learning processes for digital storytelling scientific imagination. *Eurasia Journal of Mathematics, Science and Technology Education*, *15*(2), em1659. https://doi.org/10.29333/ejmste/100636

Cremin, T., Flewitt, R., Mardell, B., & Swann, J. (Eds.). (2017). *Storytelling in early childhood: Enriching language, literacy and classroom culture*. Routledge. https://doi.org/10.4324/9781315679426

Cremin, T., Glauert, E., Craft, A., Compton, A., & Stylianidou, F. (2015). Creative little scientists: Exploring pedagogical synergies between inquiry-based and creative approaches in early years science. *Education 3-13*, *43*(4), 404–419. https://doi.org/10.1080/03004279.2015.1020655

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Eliza, D., Permana, I. A., & Dhani, M. (2025). Picture Science Story Game App: What effects on early childhood science process skills? *International Journal of Information and Education Technology*, *15*(8), 1670–1679. https://doi.org/10.18178/ijiet.2025.15.8.2368

Eshach, H., & Fried, M. N. (2005). Should science be taught in early childhood? *Journal of Science Education and Technology*, *14*(3), 315–336. https://doi.org/10.1007/s10956-005-7198-9

Esther-del-Moral-Pérez, M., López-Bouzas, N., & Castañeda-Fernández, J. (2026). Game-based storytelling with a robot character: Activating computational thinking in young children. *Early Childhood Education Journal*, *54*(3), 1293–1312. https://doi.org/10.1007/s10643-025-01930-x

Ganea, P. A., Canfield, C. F., Simons-Ghafari, K., & Chou, T. (2014). Do cavies talk? The effect of anthropomorphic picture books on children's knowledge about animals. *Frontiers in Psychology*, *5*, 283. https://doi.org/10.3389/fpsyg.2014.00283

Gelman, R., & Brenneman, K. (2004). Science learning pathways for young children. *Early Childhood Research Quarterly*, *19*(1), 150–158. https://doi.org/10.1016/j.ecresq.2004.01.009

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Annals of Family Medicine*, *13*(6), 554–561. https://doi.org/10.1370/afm.1865

Hopkins, E. J., & Weisberg, D. S. (2017). The youngest readers' dilemma: A review of children's learning from fictional sources. *Developmental Review*, *43*, 48–70. https://doi.org/10.1016/j.dr.2016.11.001

Howitt, C., Lewis, S., & Upson, E. (2011). "It's a mystery": A case study of implementing forensic science in preschool as scientific inquiry. *Australasian Journal of Early Childhood*, *36*(3), 45–55. https://doi.org/10.1177/183693911103600307

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Larkin, K., & Lowrie, T. (2022). *STEM education in the early years: Thinking about tomorrow*. Springer. https://doi.org/10.1007/978-981-19-2810-9

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

Marcus, M., Tõugu, P., Haden, C. A., & Uttal, D. H. (2021). Advancing opportunities for children's informal STEM learning transfer through parent–child narrative reflection. *Child Development*, *92*(5), e1075–e1084. https://doi.org/10.1111/cdev.13641

McClure, E. R., Guernsey, L., Clements, D. H., Bales, S. N., Nichols, J., Kendall-Taylor, N., & Levine, M. H. (2017). *STEM starts early: Grounding science, technology, engineering, and math education in early childhood*. The Joan Ganz Cooney Center at Sesame Workshop. https://joanganzcooneycenter.org/publication/stem-starts-early/

Miller-Goldwater, H. E., Hanft, M. H., Miller, A. G., & Bauer, P. J. (2024). Young children's science learning from narrative books: The role of text cohesion and caregivers' extratextual talk. *Journal of Cognition and Development*, *25*(3), 323–349. https://doi.org/10.1080/15248372.2023.2267229

Ministry of National Education [MoNE]. (2024). *Türkiye yüzyılı maarif modeli: Okul öncesi eğitim programı* [Türkiye Century Maarif Model: Preschool education program]. https://tymm.meb.gov.tr/upload/program/2024programokuloncesiOnayli.pdf

Mohana, M., Nandhini, K., & Subashini, P. (2022). Review on artificial intelligence and robots in STEAM education for early childhood development: The state-of-the-art tools and applications. In A. L. Betts &

K.-P. Thai (Eds.), *Handbook of research on innovative approaches to early childhood development and school readiness* (pp. 468–498). IGI Global. https://doi.org/10.4018/978-1-7998-8649-5.ch019

Özdemir Beceren, B., Baydemir, C., & Arı Arat, C. (2025). Exploring preschool teachers' perspectives on social and emotional learning in early childhood: A qualitative inquiry. *Journal of Pedagogical Research*, *9*(5), 180–196. https://doi.org/10.33902/JPR.202534890

Pahl, A., Fuchs, H. U., Zanella, S., & Corni, F. (2025). Interaction of physical and narrative experiencing of heat transfer phenomena by kindergarten children. *Journal of Physics: Conference Series*, *2950*(1), 012007. https://doi.org/10.1088/1742-6596/2950/1/012007

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage.

Ramanathan, G., Carter, D., & Wenner, J. (2022). A framework for scientific inquiry in preschool. *Early Childhood Education Journal*, *50*(7), 1263–1277. https://doi.org/10.1007/s10643-021-01259-1

Richert, R. A., Shawber, A. B., Hoffman, R. E., & Taylor, M. (2009). Learning from fantasy and real characters in preschool and kindergarten. *Journal of Cognition and Development*, *10*(1–2), 41–66. https://doi.org/10.1080/15248370902966594

Şahin, M. (2021). Çocuk gelişimi programı öğrencilerinin okul öncesi dönem çocuklarına yönelik STEM temelli hazırlanan etkinlikler ile ilgili görüşleri [Views of child development program students on STEM-based activities for preschool children]. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *7*(2), 518–536. https://doi.org/10.31592/aeusbed.910255

Şahin, M. (2026a). Emotional factors in inclusive early childhood education: An integrated STEM+SEL programme in a preschool context. *Acta Psychologica*, *270*, 107625. https://doi.org/10.1016/j.actpsy.2026.107625

Şahin, M. (2026b). Social sustainability of the teaching profession: Pedagogical beliefs and pre-service teachers' digital competence in STEAM. *Sustainability*, *18*(3), 1702. https://doi.org/10.3390/su18031702

Shemy, N. S. (2020). The impact of digital storytelling on motivation and achievement in teaching scientific concepts for pre-school students. *European Journal of Education Studies,* *7*(12), 801–820. https://doi.org/10.46827/ejes.v7i12.3627

Tengler, K., Kastner-Hauler, O., Sabitzer, B., & Lavicza, Z. (2022). The effect of robotics-based storytelling activities on primary school students' computational thinking. *Education Sciences*, *12*(1), 10. https://doi.org/10.3390/educsci12010010

Trundle, K. C., & Smith, M. M. (2017). A hearts-on, hands-on, minds-on model for preschool science learning. *Young Children*, *72*(1), 80–86.

Uğraş, H., Uğraş, M., Papadakis, S., & Kalogiannakis, M. (2025). Innovative early childhood STEM education with ChatGPT: Teacher perspectives. *Technology, Knowledge and Learning*, *30*(2), 809–831. https://doi.org/10.1007/s10758-024-09804-8

Uğraş, M., & Genç, Z. (2018). Investigating preschool teacher candidates' STEM teaching intention and the views about STEM education. *Bartın Üniversitesi Eğitim Fakültesi Dergisi*, *7*(2), 724–744. https://doi.org/10.14686/buefad.408150

Uyanık, G., Günşen, G., Kontaş Azaklı, T., Değirmenci, Ş., Çelik, B., Aydın, D., & Kaçira, A. B. (2026). Exploring science through stories: Investigating effects on teachers' self-efficacy and preschoolers' scientific skills and curiosity. *Early Childhood Education Journal, 54,* 5333–5346. https://doi.org/10.1007/s10643-026-02234-4

Vaknin-Nusbaum, V., & Nevo, E. (2017). A joint interactive storybook intervention program for preschool and kindergarten children. *Reading Psychology*, *38*(3), 231–261. https://doi.org/10.1080/02702711.2016.1258376

Vidal Carulla, C., Christodoulakis, N., & Adbo, K. (2021). Development of preschool children's executive functions throughout a play-based learning approach that embeds science concepts. *International Journal of Environmental Research and Public Health*, *18*(2), 588. https://doi.org/10.3390/ijerph18020588

Walker, C. M., Gopnik, A., & Ganea, P. A. (2015). Learning to learn from stories: Children's developing sensitivity to the causal structure of fictional worlds. *Child Development*, *86*(1), 310–318. https://doi.org/10.1111/cdev.12287

Wan, Z. H., Jiang, Y., & Zhan, Y. (2021). STEM education in early childhood: A review of empirical studies. *Early Education and Development*, *32*(7), 940–962. https://doi.org/10.1080/10409289.2020.1814986

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yalçın, V., & Erden, Ş. (2021). The effect of STEM activities prepared according to the design thinking model on preschool children's creativity and problem-solving skills. *Thinking Skills and Creativity*, *41*, 100864. https://doi.org/10.1016/j.tsc.2021.100864

Yang, W., Ng, D. T. K., & Su, J. (2023). The impact of story-inspired programming on preschool children's computational thinking: A multi-group experiment. *Thinking Skills and Creativity*, *47*, 101218. https://doi.org/10.1016/j.tsc.2022.101218

Zhou, W., & Yang, W. (2025). Story-based STEM design challenges in early childhood education: Child engagement and pedagogical strategies. *Early Childhood Education Journal*, *53*, 2177–2192. https://doi.org/10.1007/s10643-024-01733-6

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223. https://doi.org/10.1016/j.dr.2006.12.001

Appendix 1. Curriculum Alignment of the Activities

Table A1 maps the 12 activities onto the components of the Türkiye Century Maarif Model (TYMM) preschool program (MoNE, 2024): field skills, social–emotional learning skills (SDB), and targeted outcomes. All activities also addressed the curiosity disposition (E) and the scientific process skills recorded in the observation form.

Table A1

Alignment of the story-based STEM activities with the TYMM preschool program

3–4 years	1. Growing Chairs	Mathematics (MAB); Turkish language; Movement and health (HSAB)	Persisting with a task; following group rules	Comparing sizes; predicting whether an object fits
	2. Magical Color Bottles	Science (FAB); Art (SNAB); Turkish language	Turn-taking; sharing materials	Observing and describing color change; simple prediction
	3. Lost Keys	Science (FAB); Mathematics (MAB)	Cooperating in a small group	Testing which objects a magnet attracts; sorting by result
	4. Towers That Do Not Fall	Mathematics (MAB); Movement and health (HSAB)	Coping with failure; persisting after collapse	Building and testing stability
4–5 years	1. Tea Time in the Time Tunnel	Mathematics (MAB); Turkish language	Taking turns in role play	Sequencing events; ordering by time
	2. Entrance to the Mysterious Garden	Science (FAB); Mathematics (MAB); Turkish language	Caring for living things	Observing plant growth; comparing and measuring
	3. Opening and Closing Doors	Science (FAB); Art (SNAB); Movement and health (HSAB)	Cooperating to solve a problem	Exploring simple mechanisms; improving a design
	4. Colorful Journey	Mathematics (MAB); Movement and health (HSAB)	Giving and following directions with peers	Following and creating directional sequences (coding logic)
5–6 years	1. The Invisible Path	Science (FAB); Mathematics (MAB)	Listening to peers' ideas	Investigating light and reflection; explaining a path
	2. The Mad Hatter's Bridge	Science (FAB); Mathematics (MAB); Movement and health (HSAB)	Collaborating on a shared design; persisting after failure	Designing, testing, and strengthening a load-bearing structure
	3. The Rabbit Lost in the Dark	Science (FAB); Art (SNAB); Turkish language	Expressing and defending an idea	Relating light source, object, and shadow; explaining results
	4. The World Turned Upside Down	Science (FAB); Art (SNAB); Turkish language	Accepting different viewpoints	Observing mirror images; recording and explaining observations

Note. FAB = science field skills; MAB = mathematics field skills; HSAB = movement and health field skills; SNAB = art field skills; SDB = social–emotional learning skills; E = dispositions. Alignment is shown at the level of program components.