

The Eurasia Proceedings of Educational and Social Sciences (EPESS), 2026

Volume 52, Pages 107-119

IConSoS 2026: International Conference on Social Science Studies

Integrating Science and Sustainability in Early Childhood: Curriculum Design and Teacher-Reported Outcomes in the “Earlyscience” Program

Elif Ozturk
Giresun University

Abstract: Integrating science and sustainability in early childhood requires curriculum connections that can be examined through evidence of children’s learning. This exploratory mixed-methods study examined how the EarlyScience Program connected science content with sustainability goals and what teachers’ assessments revealed during its implementation. The 38-week program was implemented in two kindergartens in Türkiye. The matched sample included 191 children in 10 classrooms organized into 3-, 4-, 5-, and 6-year age bands. Data comprised curriculum documents, a 22-item science observation form completed by teachers in October and June, and 371 open-ended teacher notes. The initial assessment occurred approximately five weeks after implementation began. Curriculum mapping identified connections to eight Sustainable Development Goals and 14 disciplinary core ideas; seven Goals recurred across multiple months. The mean scientific attitude subscale score increased from 12.11 to 14.10 on a 5–15 scale. Mean ratings of the societal importance of science increased from 2.29 to 2.74, with the proportion rated “Always” rising from 48.2% to 74.9%. Age-band patterns were non-monotonic, and classroom means varied widely. Lexically assisted coding identified more references to science processes and engagement in the final notes, but only one note concerning nature or environmental awareness. These descriptive findings neither establish program effects nor directly demonstrate sustainability learning; ratings and notes also shared the same teacher source. The study contributes a documented example of curriculum integration and highlights the need to align assessment with sustainability aims. Future evaluations should complement broad teacher ratings with concrete evidence of children’s explanations, decisions, and actions in relevant learning contexts.

Keywords: Early childhood science, Education for sustainability, Sustainable development goals

Introduction

Young children encounter scientific phenomena and questions of shared responsibility in the same everyday settings. Watching a plant wilt can invite investigation of living things and discussion of how to care for them; exploring water can lead to questions about its properties and its use. These connections offer a basis for integrating early childhood science education with education for sustainability. Yet a shared activity does not, by itself, establish that both kinds of learning have occurred. Curriculum design must make the intended connection explicit, and assessment must provide evidence suited to each claim.

Research on early childhood environmental education documents a range of cognitive, social, emotional, and environmental outcomes, often in play-based and nature-rich settings (Ardoin & Bowers, 2020). Education for sustainability extends beyond environmental knowledge to include social and economic relationships, participation, and considered action (Hedefalk et al., 2015; Somerville & Williams, 2015). Bascopé et al. (2019) identify scientific action, community engagement, and values as complementary foundations for early sustainability education. These perspectives suggest that scientific inquiry can contribute to sustainability education, provided that inquiry is connected to questions about consequences, relationships, and choices.

The policy rationale for such connections is explicit in Target 4.7 of the 2030 Agenda, which places education for sustainable development within the knowledge and skills expected of all learners (United Nations, 2015). UNESCO (2017) distinguishes cognitive, socio-emotional, and behavioral learning objectives for the Sustainable Development Goals (SDGs). These domains are related, but evidence from one cannot automatically establish attainment in another. A child's enthusiasm during an investigation, for example, does not demonstrate understanding of resource conservation, just as naming a sustainable choice does not establish that the child makes that choice in practice.

Research Contest

Recent reviews provide a substantial basis for early science teaching while identifying uneven coverage of ages, outcomes, and research designs. O'Connor et al. (2021) describe a concentration of science concept research among children aged three to six and a smaller evidence base for younger children. Roy et al. (2025) highlight the limited representation of early childhood and children's own perspectives within scientific literacy research. Gangal et al. (2026) synthesized 17 intervention studies of early childhood science activities published between 2008 and 2023. Their findings support further investigation of science learning opportunities, but do not establish how the same activities contribute to sustainability learning.

Assessment is particularly consequential when a program is intended to serve several educational purposes. Johnstone et al. (2022) found very low certainty of evidence across the quantitative studies in their review of nature-based early childhood education and inconsistent findings for environmentally responsible behavior. These findings caution against treating a favorable general outcome as evidence for every intended benefit. Comparative studies also show that sustainability is organized differently across curricula and preschool settings (Kahrman-Pamuk & Borg, 2025; Lepičnik Vodopivec & Šindić, 2025). An explicit account of what a particular program offers, and what its measures can establish, is therefore needed.

Conceptual Framework

The EarlyScience Program treats sustainability as a possible social and environmental context for scientific inquiry. The intended connection is between understanding a phenomenon and considering why that understanding matters in everyday life. This is a curriculum proposition rather than an assumption that all science activities constitute sustainability education. A flotation investigation can address properties of materials without addressing aquatic ecosystems; learning about weather is not equivalent to understanding climate change. Sustainability interpretation requires additional framing, such as attention to resource use, living systems, or the consequences of human decisions.

The disciplinary organization draws on the Framework for K–12 Science Education (National Research Council [NRC], 2012) and the Next Generation Science Standards (NGSS Lead States, 2013). Because these standards begin at kindergarten, their use with younger children is a planning adaptation rather than a claim of formal preschool standards alignment. The focus is on accessible phenomena, investigation, and revisiting ideas through experience. The National Academies of Sciences, Engineering, and Medicine (NASEM, 2022) support opportunities for young children to engage in meaningful science and engineering, while emphasizing the importance of responsive teaching and learning environments.

Guided play provides a pedagogical rationale for combining children's exploration with adult support. Skene et al. (2022) found benefits of guided play for selected learning outcomes, with effects varying by outcome and comparison condition. Bulunuz (2013) investigated the integration of play and science instruction in kindergarten. These studies inform the program's design; they do not demonstrate that guided play mediated outcomes in the present study, because classroom interactions and implementation fidelity were not measured systematically.

A second distinction concerns the outcome of interest. The science observation form includes an item about grasping the societal importance of science. The item may be relevant to a curriculum that connects investigation with shared concerns, but it does not directly assess sustainability knowledge, values, or action. In this study it is an exploratory indicator of teachers' perceptions of the social relevance children attach to science. Neither its wording nor the form's general psychometric evidence establishes its validity as a sustainability measure. This distinction guides the interpretation of both score changes and teacher notes.

Purpose and Research Questions

The study examines how science and sustainability are connected in the EarlyScience planning documents and how teachers' observations changed during the school year. It combines descriptive longitudinal ratings, an exploratory analysis of teacher notes, and curriculum document analysis. Four questions guide the investigation:

RQ1. How are disciplinary core ideas and SDGs connected across the planned curriculum?

RQ2. What changes are observed in teacher-rated scientific attitudes, including the item on the societal importance of science?

RQ3. What descriptive patterns occur across classroom age bands and classrooms?

RQ4. How do the topics identify in teachers' open-ended notes complement or qualify the interpretation of the ratings?

Method

Research Design

The study used an exploratory convergent mixed-methods design, with quantitative ratings and qualitative notes collected at the same two measurement occasions and brought together during interpretation (Creswell & Plano Clark, 2018). Analysis of the planning documents addressed the curriculum question. The quantitative component followed one cohort without a comparison group. Accordingly, the study describes changes during program implementation and does not estimate a causal program effect. Maturation, other classroom experiences, and changes in teachers' familiarity with children remain plausible explanations for score changes (Shadish et al., 2002).

The EarlyScience Program

The "EarlyScience" Science and Nature Program was developed by the author through literature review, content mapping, and pilot activity design during the preceding academic year. Its 38-week plan spans September to June and comprises six recurring thematic modules: Life Science–Humans; Life Science and Nature–Plants and Animals; Physical Sciences; Earth and Space Science; Matter Science; and Current Issues. These modules are distributed across ten monthly planning periods.

The program begins with direct investigation of familiar phenomena and uses accessible materials, including leaves, soil, worms, water, sponges, ramps, rocks, and ice. Concepts recur in different contexts across the year. For example, investigations of animal movement are revisited through comparisons of animal coverings and life cycles. Each week provides parallel activity specifications for the 3–4 and 5–6 age bands. In the human body module, younger children explore the heart, hands, and feet, whereas older children also investigate digestion and smell. Fifteen units draw on the Great Explorations in Math and Science materials developed at the Lawrence Hall of Science (Barrett et al., 1999).

The design-stage documents associate activities with NGSS disciplinary core ideas and selected SDGs, with Target 4.7 used as a cross-cutting educational frame. The mapping records intended curricular connections. It does not establish that every associated sustainability idea was explicitly discussed in each classroom, nor does the number of weeks assigned to a Goal indicate the amount of time each child spent on that Goal. Teachers delivered the program during regular classroom time; no systematic measure of dosage or implementation fidelity was collected.

Setting and Participants

Two kindergartens in İstanbul, Türkiye, implemented the program during the 2025–2026 academic year. School A contributed eight classrooms spanning all four age bands, and School B contributed two classrooms in the 5- and 6-year bands. Of 196 children assessed at the first occasion and 209 at the second, 191 could be matched across occasions and constitute the analytic sample. Five children assessed initially had left before the final

assessment, and 18 children joined during the year. These unmatched cases were excluded from the longitudinal ratings.

The matched sample included 34 children in the 3-year band, 30 in the 4-year band, 65 in the 5-year band, and 62 in the 6-year band, across 10 classrooms. Age refers to the classroom age band at the first measurement, not to exact chronological age. Age in months, socioeconomic circumstances, and language background were not available. The participating settings therefore provide a contextual sample rather than a representative sample of preschool children in Türkiye.

Data Sources

Science Observation Form

Teachers completed the Science Observation Form, whose Turkish-language validity and reliability were examined by Kuru and Akman (2017). The form was used with written permission obtained from Kuru and Akman. Their article reports a three-factor, 22-item structure and a validation analysis involving 370 children aged 48–66 months, with RMSEA = .04, CFI = .98, and Cronbach's $\alpha = .93$ for the full form. The validation sample is distinct from the 250-child sample used for the substantive analyses in that article. The source describes a Turkish version of an original form; it is therefore cited here for its Turkish psychometric evidence rather than as proof of original instrument development. The administered form contains 14 science process items, three problem-solving items, and five scientific attitude items. Items were coded Never (1), Sometimes (2), and Always (3). The scientific attitude subscale ranges from 5 to 15 and addresses confidence, flexibility, persistence, awareness of one's learning, and the societal importance of science. Internal consistency in the present sample was $\alpha = .90$ at the first measurement and $\alpha = .86$ at the second. These estimates concern consistency among items, not the validity of a single item as a sustainability indicator. The societal-importance item receives particular attention because of its conceptual relationship to the program's social framing of science. Its use remains exploratory. The Turkish validation evidence covers 48–66 months, whereas the present classroom bands extend below and potentially above that range. Validity across the full age span, longitudinal measurement invariance, and equivalence between age bands were not established. The English item labels used in this article describe the administered Turkish items and do not constitute a separately validated English version.

Teacher Observation Notes

At each measurement occasion, teachers were invited to record what stood out about each child during science activities. The prompt did not explicitly request observations of sustainability. The matched sample yielded 371 nonempty notes: 181 at the first occasion and 190 at the second. Thus, 10 initial notes and one final note were missing. Notes and ratings were produced by the same teachers and are treated as complementary accounts from a shared source, not as independent validation of one another.

Curriculum Planning Documents

The document source comprised the design-stage plans specifying weekly themes, age-differentiated activities, disciplinary core ideas, and SDG associations. These documents were prepared before implementation. Analysis summarized the program designer's intended connections rather than independently validating the mapping or auditing its enactment.

Data Collection and Ethics

Teachers completed the initial ratings on October 11–12, 2025, approximately five weeks after the program began, and the final ratings between June 15 and 25, 2026. The interval was approximately eight months. The October assessment is therefore an early-program measurement rather than a pre-intervention baseline. Calendar time between measurements should not be equated with instructional weeks or individual exposure. The same class teachers delivered the activities and provided the ratings and notes. Institutional permission was obtained from both kindergartens, and written informed consent was obtained from the parents or legal guardians of participating children. Child names were replaced with codes before analysis. Teacher observations are

described through paraphrases rather than identifying quotations. The author's roles as program developer and investigator are acknowledged in interpreting the study.

Data Analysis

Curriculum coverage was summarized in a matrix of monthly planning periods, SDGs, and disciplinary core ideas. The unit of coverage was the planned module period, not an observed child-level learning episode. The number of distinct Goals and core ideas, recurrence across months, and nominal weeks associated with each Goal were calculated from the mapping. Weeks could contribute to more than one Goal and were not added to produce a total exposure score.

The quantitative presentation is descriptive. Means, standard deviations, mean changes, and the proportion rated Always summarize the scientific attitude items and subscale. Mean changes use unrounded values, so subtraction of the displayed rounded means may differ by .01. For the subscale, Cohen's d_z summarizes the paired change as the mean difference divided by the standard deviation of individual differences (Fritz et al., 2012). It is an unadjusted within-sample index, not a treatment effect. Age-band summaries and the range of classroom means describe variation without attributing it to development or rater differences. Spearman correlations describe the association of societal-importance ratings and changes with the corresponding mean across the 14 science process items.

Three-point items were summarized numerically for readability, but their ordered categories should not be assumed to be equally spaced. The proportion rated Always is consequently reported alongside means and is not labeled mastery, because no independently validated mastery threshold was available. Ceiling-constrained scores, clustered observations, and the absence of a comparison group limit interpretation. Population-level significance tests, claims of equivalent gains across different outcomes, and adjusted developmental effects are not advanced in this report.

Teacher notes underwent exploratory, lexically assisted content analysis, drawing on the directed approach described by Hsieh and Shannon (2005). Seven categories were derived from the form's constructs and the sustainability aims and refined during an initial reading: higher-order process skills; observation and inquiry practice; willing engagement; curiosity and question asking; confidence and persistence; nature and environmental awareness; and transfer to everyday life. Turkish lexical stems were applied to case-normalized text. Each detected match was inspected in context and false positives were removed. This was necessary, for example, when a waste-related stem appeared within the Turkish word for mathematics or a recycling-related stem appeared within an idiom meaning to respond.

The unit counted was a note containing at least one verified match for a category; a note could contribute to several categories. Percentages use the number of available notes at each occasion, 181 and 190, rather than assigning missing notes an absence code. These frequencies describe the observed texts and are not paired estimates of change in children's behavior. Coding was conducted by one analyst, and no intercoder reliability estimate was available. Checking detected matches reduces false positives but does not establish that all relevant meanings were detected among unmatched texts. Category absence is therefore interpreted as absence of identified wording, not absence of the underlying learning or behavior.

Integration occurred through comparison of the intended curriculum connections, rating patterns, and topics identified in the notes. Convergence was considered at the level of broad themes, and discrepancies were examined in relation to the different prompts, constructs, and observation formats. Neither stream was treated as a criterion measure for the other.

Results and Discussion

Planned Curriculum Connections

The mapping associated the program with eight SDGs: good health and well-being; clean water and sanitation; affordable and clean energy; sustainable cities and communities; responsible consumption and production; climate action; life below water; and life on land. Fourteen distinct disciplinary core idea codes were represented: five in life science, four in physical science, and five in Earth and space science. Table 1

summarizes the planned associations. Target 4.7 functions as the cross-cutting frame and is not counted as a ninth content Goal.

Table 1. Planned SDG and science connections

SDG	Modules and months	Representative investigations	Core idea codes	Weeks
3 Good health and well-being	Life Science–Humans (Sep); Physical Sciences (Nov)	Our heart; what happens to what I eat; stop the germ!; we need healthy teeth; interesting medicines	LS1.A; LS1.C	8
6 Clean water and sanitation	Physical Sciences (Jan); Matter Science (Feb)	Sponges and water; melting ice; properties of soap and water	PS1.A; PS1.B	8
7 Affordable and clean energy	Matter Science (Feb)	Batteries and electricity; electricity in our lives; different ways of producing light	PS3.B; PS4.B	4
11 Sustainable cities and communities	Earth and Space (Dec); Life Science and Nature (May)	Litter hunt; we can care for the Earth	ESS3.A; ESS3.C	8
12 Responsible consumption and production	Earth and Space (Dec); Life Science and Nature (Mar, May)	Litter hunt; products people obtain from animals and plants; life cycles	ESS3.C; LS1.B	12
13 Climate action	Earth and Space (Dec, Apr); Life Science and Nature (May); Current Issues (Jun)	Rain and snow; dressing for the season; types of cloud; who has seen the wind?; climate change	ESS2.D; ESS3.B; ESS3.C	14
14 Life below water	Matter Science (Feb); Life Science and Nature (Mar)	Sinking or floating a boat; exploring fish	PS1.A; LS4.D	8
15 Life on land	Life Science and Nature (Oct, Mar, May); Physical Sciences (Jan); Earth and Space (Apr)	Observing a worm; comparing leaves; digging soil; animal coverings; classifying rocks; wild and domestic animals	LS2.A; LS4.D; ESS2.A	20

SDG = Sustainable Development Goal. Codes follow NGSS Lead States (2013). Weeks denote nominal durations of associated planning periods; they do not measure observed instructional dosage. Connections are the program designer’s intended associations. A week may be associated with multiple Goals. Life science and nature concerns plants and animals.

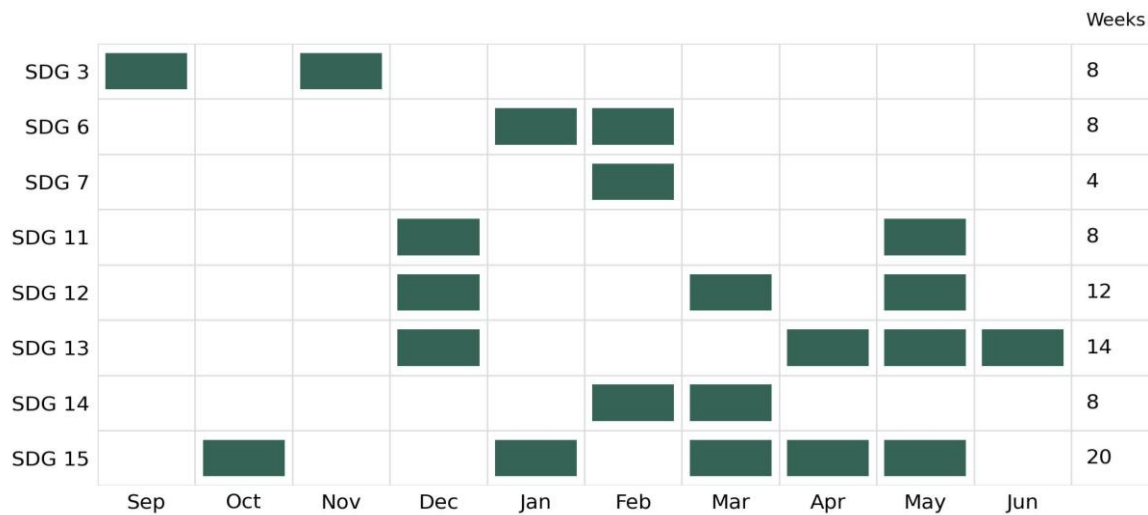


Figure 1. SDG recurrence across monthly plans

Seven of the eight Goals appeared in at least two monthly periods; affordable and clean energy appeared in one. The mean was 2.6 months per Goal. Every month included at least one Goal, with a range of one to four (Figure

1). Life on land was associated with 20 nominal weeks and affordable and clean energy with four. These figures describe the breadth and recurrence of the plan. They do not establish the depth of sustainability discussion, the fidelity of implementation, or the achievement of an SDG-related learning objective.

Filled cells represent intended monthly connections. The right-hand column shows nominal weeks associated with each Goal. The chart summarizes planning coverage, not verified exposure or learning. This mapping offers a concrete way to organize disciplinary content alongside sustainability intentions. It is consistent with the attention to scientific action and context in Bascopé et al. (2019), and with the different curricular arrangements discussed by Kahrman-Pamuk and Borg (2025). However, some links are closer to the underlying sustainability concern than others. Caring for the Earth and investigating litter explicitly invite attention to human impacts; flotation, electricity, or weather require additional pedagogical framing to connect them with aquatic environments, clean energy, or climate action. The present mapping identifies these intended opportunities rather than demonstrating that the connections were experienced by children.

Changes in Scientific Attitude Ratings

The scientific attitude subscale mean increased from 12.11 ($SD = 2.72$) in October to 14.10 ($SD = 1.57$) in June, a mean increase of 1.99 points on the 5–15 scale. The unadjusted standardized paired change was $d_z = 0.82$. All five-item means increased (Table 2). The proportion at the maximum subscale rose from 27.7% to 66.5%, indicating substantial ceiling concentration at the final measurement.

Table 2. Scientific attitude ratings across occasions

Item	October M (SD)	June M (SD)	Mean change	Always October %	Always June %
Shows confidence in science activities	2.49 (0.59)	2.87 (0.33)	+0.39	53.4	87.4
Shows flexibility	2.36 (0.62)	2.87 (0.35)	+0.51	43.5	87.4
Shows persistence in scientific work	2.48 (0.61)	2.86 (0.36)	+0.38	53.9	86.9
Becomes aware of own learning	2.49 (0.62)	2.75 (0.44)	+0.26	56.0	75.9
Grasps the societal importance of science	2.29 (0.77)	2.74 (0.45)	+0.46	48.2	74.9
Subscale total (range 5–15)	12.11 (2.72)	14.10 (1.57)	+1.99	27.7	66.5

$N = 191$ matched children. Items are coded Never = 1, Sometimes = 2, and Always = 3. For the subscale total, percentages denote the maximum score of 15. Changes use unrounded values. Always is a response category, not a validated mastery threshold. Results describe the sample and are not adjusted program effects.

The societal-importance item increased from 2.29 ($SD = 0.77$) to 2.74 ($SD = 0.45$). The proportion rated Always increased from 48.2% to 74.9%, a difference of 26.7 percentage points. Teachers therefore gave more favorable ratings of this aspect of children's engagement with science by June. The observation does not establish that children acquired a distinct sustainability competence, because that interpretation was not validated for the item. Changes on the societal-importance item were positively associated with changes in the mean of the 14 science process items (Spearman's $\rho = .64$); the corresponding association between final scores was $\rho = .65$. These are descriptive pooled correlations. They do not distinguish within-classroom relationships from differences between classrooms and may reflect shared development, common classroom experiences, or common-method influences. They neither establish that the item measures a distinct construct nor demonstrate that it lacks one.

The increase is compatible with the possibility that sustained opportunities for science supported children's participation and teachers' recognition of their capabilities. It is also compatible with maturation, broader classroom learning, and changing teacher expectations. Research on science through play and guided play provides a rationale for such opportunities (Bulunuz, 2013; Skene et al., 2022), but cannot resolve these alternatives in this cohort. The appropriate conclusion is that teacher-rated scientific attitudes increased during the observed period. Comparisons of improvement across different constructs would require attention to baseline distributions, measurement properties, and the large ceiling effects.

Age Bands and Classroom Context

Table 3 shows the societal-importance ratings by classroom age band. Means increased in each band, but the cross-sectional pattern was non-monotonic. The 4-year band had the lowest mean at both occasions, while the 5-year band began close to the maximum. Three-year-olds were rated higher than four-year-olds in October and June. These patterns do not support simple developmental ordering.

Table 3. Societal-importance ratings by classroom age band

Age band	n	October M (SD)	June M (SD)	Mean change	Always October %	Always June %
3 years	34	2.18 (0.63)	2.65 (0.49)	+0.47	29.4	64.7
4 years	30	1.10 (0.31)	2.07 (0.25)	+0.97	0.0	6.7
5 years	65	2.77 (0.49)	2.95 (0.21)	+0.18	80.0	95.4
6 years	62	2.42 (0.62)	2.90 (0.35)	+0.48	48.4	91.9

Age bands refer to classrooms at the initial measurement, not exact age in months. Ratings range from 1 to 3. Changes use unrounded values. Age bands differ in classroom representation and initial score distributions; the table is descriptive.

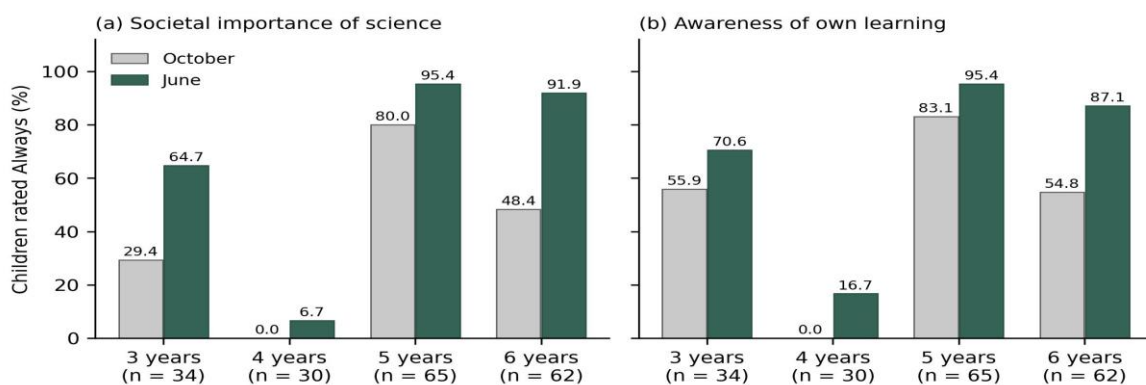


Figure 2. Highest-category ratings by classroom age band

Among four-year-olds, the mean increased from 1.10 to 2.07, while the proportion rated Always increased from 0% to 6.7%. Among five-year-olds, the corresponding proportions were 80.0% and 95.4%. A similar contrast appeared on awareness of one's learning (Figure 2). A low percentage in the highest category is not equivalent to absence of emerging competence, particularly where middle-category responses predominate. Percentages refer to children rated Always. Panels show societal importance of science and awareness of one's learning. These distributions do not establish developmental thresholds.

Classroom means also differed substantially. Initial scientific attitude subscale means ranged from 8.20 to 14.47 on the 5–15 scale, and initial societal-importance means ranged from 1.07 to 2.95 on the 1–3 scale. These ranges describe classroom variation without identifying its source. Classroom composition, prior experiences, learning opportunities, and teachers' interpretations of the rating categories may all contribute. Because teachers assessed their own classrooms, teacher and classroom influences cannot be separated.

Waterman et al. (2012) demonstrate that assessor variation can be consequential in early childhood assessment, making it a plausible concern here rather than an established explanation for these differences. Likewise, observations of metacognition and self-regulation among three- to five-year-olds show the importance of attending to concrete verbal and nonverbal indicators (Whitebread et al., 2009). The present results therefore do not justify withholding sustainability discussion from younger children. They support using accessible examples and gathering evidence that is sensitive to children's different ways of expressing understanding.

What Teachers Recorded in Their Notes

Table 4 reports the frequency of verified lexical matches in available notes. References to higher-order process skills appeared in two of 181 October notes and 20 of 190 June notes. Observation and inquiry practice appeared in 25 and 52 notes, respectively, and willing engagement in 53 and 79. These descriptive distributions suggest that teachers' written accounts gave greater prominence to process skills and participation at the final occasion. They do not establish a statistically tested change in the behavior of matched children.

Table 4. Topics identified in available teacher notes

Code	October n	October %	June n	June %
Higher-order process skills	2	1.1	20	10.5
Observation and inquiry practice	25	13.8	52	27.4
Willing engagement	53	29.3	79	41.6
Curiosity and question asking	29	16.0	33	17.4
Confidence and persistence	3	1.7	3	1.6
Nature and environmental awareness	0	0.0	1	0.5
Transfer to everyday life	0	0.0	1	0.5

Denominators are 181 nonempty October notes and 190 nonempty June notes. Ten initial notes and one final note were missing. Each note may receive several codes, so percentages need not sum to 100. Counts concern verified lexical matches, not the prevalence of children's competencies. Missing notes is excluded from these percentages.

Only one note contained a verified match for nature and environmental awareness, and one for transfer to everyday life; both were recorded in June. In paraphrase, the first described awareness of protecting nature and animals, and the second described applying learning in everyday life. The latter does not necessarily concern sustainability. The counts identify what was explicitly captured by the coding procedure and should not be treated as an exhaustive inventory of children's sustainability-related actions or ideas.

The juxtaposition is informative: 74.9% of children were rated Always on the societal-importance item in June, yet explicit nature-related language was rare in the accompanying notes. Several explanations remain possible. The rating item and the environmental code concern different constructs; teachers may have focused their notes on participation or science processes; brief notes may omit events that informed a rating; and the lexical procedure may miss relevant meanings. Because the prompt did not ask about sustainability, the notes cannot function as a sensitive test of its presence or absence.

The qualitative contribution is consequently a qualification of interpretation. The notes provide examples and topic distributions that are broadly compatible with increasingly favorable science ratings, but they offer little explicit detail about sustainability. This gap prevents the ratings from being interpreted as independently corroborated sustainability learning. It does not demonstrate that the scale is invalid for its intended science constructs or that sustainability learning failed to occur. The need to distinguish these conclusions is consistent with the conceptual breadth described in reviews of early sustainability education (Hedefalk et al., 2015; Somerville & Williams, 2015).

Integrating Curriculum and Assessment Evidence

Across the three sources, the most defensible finding is a difference between the specificity of the curriculum intentions and the specificity of the outcome evidence. Planning documents provide repeated opportunities to connect science with selected SDGs. Teacher ratings document more favorable judgments of scientific attitudes over time. The notes provide a complementary, selectively recorded account dominated by science processes and engagement. None of these sources independently establishes how children understood or acted on the intended sustainability connections.

This distinction matters when adapting a curriculum to different settings. The program's contribution is a documented example of how connections can be planned across a school year, together with evidence about what routine teacher assessment makes visible and leaves unresolved. The study does not show that the integrated design outperforms science-only provision, saves instructional time, or develops each mapped sustainability outcome. Teacher-rated outcomes in eco-program research also need to be read in relation to their measurement source (Lepičnik Vodopivec & Šindić, 2025), while the broader review evidence cautions against generalizing from one outcome domain to another (Ardoin & Bowers, 2020; Johnstone et al., 2022).

Limitations

The absence of a comparison group prevents separation of program-associated change from maturation and other experiences. The initial measurement took place after approximately five weeks of implementation, so it

does not represent the children's status before the program. Teachers were both implementers and assessors, and the researcher was the program developer. These overlapping roles may influence expectations and interpretation. The program mapping was not independently audited, and fidelity, instructional dosage, and the content of sustainability discussions were not systematically observed.

The sample is limited to two settings and 10 classrooms. Exact age and family-background information were unavailable, and classroom age bands are partly confounded with setting and teacher. The instrument's published Turkish validation range does not cover the entire range represented by these classroom bands. Measurement invariance across time and age groups was not assessed. Three-category ratings, pronounced ceiling effects, and the absence of a validated sustainability outcome restrict what numerical changes can mean. The descriptive results and standardized paired change do not provide cluster-adjusted estimates or uncertainty intervals for a wider population.

The notes were short, unstructured accounts from the same teachers who provided the ratings. They were not prompted specifically for sustainability, and some notes were missing. Lexical coding by one analyst can overlook relevant meanings, and no intercoder reliability estimate or independent assessment was available. Differences between the two note distributions therefore concern the recorded corpus. These limitations prevent a conclusion about the prevalence or absence of sustainability learning, but they clarify the additional evidence required to investigate it.

Conclusion

This study provides a practice-based example of how science and sustainability can be connected within a year-long early childhood curriculum. The EarlyScience Program's planning documents linked scientific content to selected Sustainable Development Goals through recurring opportunities across the school year. In the matched cohort, teachers reported more favorable scientific attitudes at the final measurement, including greater recognition of the societal importance of science. Their written notes also gave greater prominence to inquiry, observation, and participation. Together, these findings document a curriculum approach and patterns in teachers' observations that can inform the design and assessment of integrated early childhood learning experiences.

The study's central contribution is to clarify the relationship between curriculum intentions and the evidence used to assess them. Planned sustainability connections, positive science ratings, and descriptions of children's understanding offer distinct perspectives on the educational process. Although the observed changes cannot be attributed solely to the program or interpreted as direct evidence of sustainability learning, examining these sources together reveals where curriculum and assessment need closer alignment. In particular, broad ratings of scientific attitudes should be complemented by concrete accounts of how children explain, question, and respond to environmental and social concerns.

For early childhood education, this points toward an approach in which science-sustainability connections are made explicit in both learning experiences and documentation. Everyday encounters with water, materials, living things, and shared environments can provide contexts for such connections. The next step is to capture how children make sense of these experiences through their questions, explanations, choices, and actions. The EarlyScience Program offers a concrete starting point for this work: a documented curriculum structure that future research can refine and evaluate using evidence more closely aligned with its sustainability aims.

Recommendations

Curriculum developers should specify how each intended SDG connection becomes available to children: the phenomenon investigated, the question connecting it with a shared concern, and the explanation or action that could make learning observable. Activity labels alone are insufficient. Teachers can invite age-responsive discussion of everyday situations, including caring for living things, using water, sharing materials, and considering waste. The present age-band patterns do not support a fixed age threshold for such discussion.

Assessment should combine concrete examples supporting each rating with brief observations of relevant situations and opportunities for children to explain their choices. Prompts should allow verbal, practical, and nonverbal responses. Whitebread et al. (2009) illustrate the value of observable indicators in assessing young children, while Paños et al. (2022) show that specific questioning skills can be supported through instruction.

These sources inform assessment design without establishing the validity of any new sustainability indicator. Proposed indicators require piloting and evidence of reliability and construct relevance.

Future research should include comparison conditions, an assessment before implementation begins, exact age in months, implementation records, and independent observations where feasible. Longitudinal analyses should account for repeated observations within children and clustering within classrooms. Studies of sustainability should use outcomes that directly address the intended knowledge, values, and actions, rather than relying on a single general science-attitude item. Such evidence would allow the curricular proposition examined here to be tested more directly.

Scientific Ethics Declaration

* The author declares that the scientific ethical and legal responsibility of this article published in EPESS Journal belongs to the author.

* Institutional permission was obtained from both participating kindergartens, and written informed consent was obtained from the parents or legal guardians of participating children. Child identities were coded before analysis

Conflict of Interest

* The author developed the EarlyScience Program and conducted this study. This professional involvement is disclosed as relevant to interpretation of the findings. The author declares no other conflict of interest.

Funding

* This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgements or Notes

* This article was presented as an oral presentation at the International Conference on Social Science Studies (www.iconsos.net) held in İstanbul/Türkiye on September 17-18, 2026

* The author thanks the teachers of the participating classrooms, the children and their families, and Berrin Akman and Nilüfer Kuru for permission to use the Science Observation Form.

* An AI assistant supported literature searching, language and structural editing, and preparation of figures from the numerical values reported in the manuscript. The author is responsible for the study design, data, analyses, interpretation, and final content.

References

- Ardoin, N. M., & Bowers, A. W. (2020). Early childhood environmental education: A systematic review of the research literature. *Educational Research Review*, 31, Article 100353. <https://doi.org/10.1016/j.edurev.2020.100353>
- Barrett, K., Blinderman, E., Boffen, B., Echols, J., House, P. A., Hosoume, K., & Kopp, J. (1999). *Science and math explorations for young children: A GEMS/PEACHES handbook for early childhood educators, childcare providers, and parents*. Lawrence Hall of Science, University of California, Berkeley. <https://eric.ed.gov/?id=ED437165>
- Bascopé, M., Perasso, P., & Reiss, K. (2019). Systematic review of education for sustainable development at an early stage: Cornerstones and pedagogical approaches for teacher professional development. *Sustainability*, 11(3), Article 719. <https://doi.org/10.3390/su11030719>

- Bulunuz, M. (2013). Teaching science through play in kindergarten: Does integrated play and science instruction build understanding? *European Early Childhood Education Research Journal*, 21(2), 226–249. <https://doi.org/10.1080/1350293X.2013.789195>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2–18. <https://doi.org/10.1037/a0024338>
- Gangal, M., Çalık, M., & Ozturk, Y. (2026). Effectiveness of early childhood science activities in promoting scientific literacy: A meta-analysis study. *Science & Education*. Advance online publication. <https://doi.org/10.1007/s11191-026-00748-2>
- Hedefalk, M., Almqvist, J., & Östman, L. (2015). Education for sustainable development in early childhood education: A review of the research literature. *Environmental Education Research*, 21(7), 975–990. <https://doi.org/10.1080/13504622.2014.971716>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Johnstone, A., Martin, A., Cordovil, R., Fjørtoft, I., Iivonen, S., Jidovtseff, B., Lopes, F., Reilly, J. J., Thomson, H., Wells, V., & McCrorie, P. (2022). Nature-based early childhood education and children's social, emotional and cognitive development: A mixed-methods systematic review. *International Journal of Environmental Research and Public Health*, 19(10), Article 5967. <https://doi.org/10.3390/ijerph19105967>
- Kahriman-Pamuk, D., & Borg, F. (2025). Sustainability, play and children's agency: A comparative study of the Swedish and Turkish preschool curricula. *European Early Childhood Education Research Journal*, 33(4), 659–675. <https://doi.org/10.1080/1350293X.2024.2415948>
- Kuru, N., & Akman, B. (2017). Examining the science process skills of preschoolers with regards to teachers' and children' variables. *Education and Science*, 42(190), 269–279. <https://doi.org/10.15390/EB.2017.6433>
- Lepičnik Vodopivec, J., & Šindić, A. (2025). Early childhood education for sustainability in preschools with eco-programs: Implementation and learning outcomes. *Journal of Baltic Science Education*, 24(1), 133–148. <https://doi.org/10.33225/jbse/25.24.133>
- National Academies of Sciences, Engineering, and Medicine. (2022). *Science and engineering in preschool through elementary grades: The brilliance of children and the strengths of educators*. The National Academies Press. <https://doi.org/10.17226/26215>
- National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. <https://doi.org/10.17226/13165>
- NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press. <https://doi.org/10.17226/18290>
- O'Connor, G., Fragkiadaki, G., Fleer, M., & Rai, P. (2021). Early childhood science education from 0 to 6: A literature review. *Education Sciences*, 11(4), Article 178. <https://doi.org/10.3390/educsci11040178>
- Paños, E., Carrión, A., & Ruiz-Gallardo, J.-R. (2022). Promoting questioning in early childhood science education. *International Journal of Science Education*, 44(11), 1840–1854. <https://doi.org/10.1080/09500693.2022.2099593>
- Roy, G., Sikder, S., & Danaia, L. (2025). Adopting scientific literacy in early years from empirical studies on formal education: A systematic review of the literature. *International Journal of STEM Education*, 12, Article 26. <https://doi.org/10.1186/s40594-025-00547-1>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development*, 93(4), 1162–1180. <https://doi.org/10.1111/cdev.13730>
- Somerville, M., & Williams, C. (2015). Sustainability education in early childhood: An updated review of research in the field. *Contemporary Issues in Early Childhood*, 16(2), 102–117. <https://doi.org/10.1177/1463949115585658>
- UNESCO. (2017). *Education for Sustainable Development Goals: Learning objectives*. <https://unesdoc.unesco.org/ark:/48223/pf0000247444>
- United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development* (A/RES/70/1). <https://sdgs.un.org/2030agenda>

- Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2012). The matter of assessor variance in early childhood education—Or whose score is it anyway? *Early Childhood Research Quarterly*, 27(1), 46–54. <https://doi.org/10.1016/j.ecresq.2011.06.003>
- Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning*, 4(1), 63–85. <https://doi.org/10.1007/s11409-008-9033-1>

Author(s) Information

Elif Ozturk

Giresun University, Faculty of Education, Department of
Early Childhood Education, Giresun, Türkiye
Visiting Research Scholar, Indiana University Bloomington,
School of Education, Department of Curriculum and
Instruction, Bloomington, Indiana, USA
ORCID iD: <https://orcid.org/0000-0002-3764-4526>

*Contact e-mail: elif.ozturk@giresun.edu.tr

*Corresponding author(s) contact email

To cite this article:

Ozturk, E. (2026). Integrating science and sustainability in early childhood: Curriculum design and teacher-reported outcomes in the “Earlyscience” program. *The Eurasia Proceedings of Educational and Social Sciences (EPESS)*, 52, 107-119.