Research Article
Araştırma Makalesi

https://doi.org/10.21764/maeuefd.1918185

Early Childhood Education for Sustainable Development: Development of a Checklist Based on the 7R Framework

Sürdürülebilir Kalkınma İçin Erken Çocukluk Eğitimi: 7R Çerçevesine Dayalı Bir Kontrol Listesinin Geliştirilmesi

Gülcan Gül TOKER [1*] **Fatma ALİSİNANOĞLU** [2] **Simge YILMAZ UYSAL** [3]

[1] Fatih Sultan Mehmet Foundation University, Graduate Education Institute, Basic Education Department, İstanbul, Türkiye
[1] METU Kindergarten, Directorate of Health, Culture & Sports, METU, Ankara, Türkiye
[2] Gazi University, College of Education, Department of Primary Education, Ankara, Türkiye
[3] Mersin University, College of Education, Department of Primary Education, Mersin, Türkiye

Abstract: The aim of this study is to develop an observation-based checklist to assess young children's (60–72 months) sustainable development behaviors based on the 7R framework proposed by the OMEP (Respect, Rethink, Reflect, Recycle, Redistribute, Reduce, Reuse) and to provide initial evidence of its structural validity. The study employed quantitative research design, and the theoretically defined seven-dimensional structure was tested using confirmatory factor analysis (CFA). Data were collected from 448 children evaluated through preschool teachers' observations across Türkiye. The instrument, initially consisting of 51 items, was reduced to 35 items. Second-order CFA results indicated acceptable model fit (χ^2/df = 3.03; RMSEA = .067; SRMR = .04) and supported the proposed structure.

Keywords: Education for sustainable development, 7R framework, observation checklist, confirmatory factor analysis

Özet: Bu çalışmanın amacı, OMEP tarafından önerilen 7R çerçevesi (Saygı Duy, Sorgula, Yansıt, Geri Dönüştür, Yeniden Dağıt, Azalt, Tekrar Kullan) temelinde, erken çocukluk döneminde (60–72 ay) çocukların sürdürülebilir kalkınmaya yönelik davranışlarını değerlendirmek amacıyla gözleme dayalı bir kontrol listesi geliştirmek ve aracın yapı geçerliğine ilişkin ilk kanıtları sunmaktır. Araştırma nicel araştırma deseni ile yürütülmüş ve kuramsal olarak tanımlanan yedi boyutlu yapı doğrulayıcı faktör analizi (DFA) ile test edilmiştir. Veriler, Türkiye genelindeki okul öncesi kurumlarda görev yapan öğretmenlerin gözlemlerine dayalı olarak değerlendirilen 448 çocuktan elde edilmiştir. Başlangıçta 51 maddeden oluşan araç, aşamalı revizyon süreci sonucunda 35 maddeye indirgenmiş ve ikinci düzey DFA sonuçları model uyumunun kabul edilebilir olduğunu göstermiştir.

Anahtar Kelimeler: Sürdürülebilir kalkınma için eğitim, 7R çerçevesi, gözlem kontrol listesi, doğrulayıcı faktör analizi

1. Introduction

Sustainable development is defined as a holistic approach that aims to meet the needs of the present without jeopardizing the ability of future generations to meet their own needs, and includes environmental, economic, and socio-cultural dimensions (World Commission on Environment and Development [WCED], 1987; Purvis et al., 2019). Davis (2009) emphasizes that education plays a critical role in achieving the Sustainable Development Goals. Accordingly, Early Childhood Education for Sustainability (ECEfS) is regarded as important for supporting the development of behaviors consistent with the Sustainable Development Goals from the early years of life (UNESCO, 2017). Pramling Samuelsson & Kaga (2008) state that early childhood is a fundamental developmental stage in which

*Corresponding author: gultoker@metu.edu.tr
Received: *28/03/2026*
Revised: *23/07/2026*
Accepted: *23/07/2026*

© Author(s) 2026. This work is distributed under https://creativecommons.org/licenses/by/4.0/

values, attitudes, and behavioral patterns are shaped, and therefore, education for sustainable development should begin as early as possible. It is emphasized that these values and behavioral patterns acquired at an early age form an important foundation for the development of environmental and social responsibilities in individuals' later lives (Davis, 2009; Engdahl, 2015). For example, Davis (2009) states that sustainability-based educational practices implemented during early childhood strengthen children's connection with nature and support the development of environmental responsibility. In a similar study, it is noted that sustainability-focused learning processes in early childhood education settings support children in developing a sense of responsibility and care for the environment (Ärlemalm-Hagsér, 2012). It is emphasized that children at this stage not only develop environmental awareness but also learn behaviors such as social responsibility, sharing, and resource management through experience (Güler Yıldız et al., 2021). However, there is a significant body of research indicating that sustainability-based learning experiences in early childhood support children's social-emotional development, critical thinking skills, and environmental awareness (Chapman & O'Gorman, 2022; Cutter-Mackenzie & Edwards, 2013; Davis & Elliott, 2023; Engdahl, 2015; Güler Yıldız et al., 2021; Kahriman-Pamuk & Pramling Samuelsson, 2024; Köklü-Yaylacı & Feriver, 2020). A review of the existing research shows that young children are capable of thinking about environmental issues and can meaningfully participate in sustainability initiatives (Borg & Pramling Samuelsson, 2022; Engdahl, 2015; Engdahl & Rabušicová, 2011).

The 7R principles proposed by the World Organisation for Early Childhood Education (OMEP) for translating these findings into a concrete and pedagogically applicable framework in the context of early childhood -Respect, Rethink, Reflect, Recycle, Redistribute, Reduce, and Reuse-provide an important roadmap (UNESCO, 2017). Borg & Pramling Samuelsson (2022) note that abstract sustainability goals take on meaning at the early childhood level through the 7R approach. ECEfS is not limited to the mere transfer of information. It also serves as a holistic pedagogical framework that supports children's processes of thinking, questioning, reflecting, and taking action (Siraj-Blatchford & Pramling Samuelsson, 2016).

It is well known that assessment tools are essential for systematically tracking learning outcomes in ECEfS (UNESCO, 2017). Van Harskamp et al., (2025) note that there is a growing need for assessment tools to effectively implement and develop sustainability education in early childhood. Therefore, there is a significant need to assess the learning outcomes of ECEfS using reliable and valid measurement tools (Davis & Elliott, 2023; Siraj-Blatchford & Pramling Samuelsson, 2016).

Various assessment tools have been developed in the literature to evaluate attitudes, knowledge, and behaviors related to sustainability during early childhood. Among these tools are the Children's Environmental Attitudes and Knowledge Scale (CHEAKS) (Leeming et al., 1995), the Children's Attitudes Toward the Environment Scale-Preschool Version (CATES-PV) (Musser & Diamond, 1999), and the Children's Environmental Attitudes and Behaviors Scale (Evans et al., 2007). In addition, the Environmental Awareness Assessment Scale was developed to determine children's levels of environmental awareness (Gökçeli & Kandır, 2016). Similarly, the Environmental Sustainability Scale for Children (60–72 months) was developed to assess the awareness and understanding levels of preschool children regarding sustainable environmental behaviors (Özkan et al., 2020).

Tools for evaluating sustainability practices in early childhood education settings are also available in the literature. The Early Childhood Environmental Education Rating Scale (ECEERS), developed for this purpose, focuses on assessing the quality of environmental education practices implemented in early childhood settings (Bhagwanji, 2011). In addition,

the Environmental Rating Scale for Sustainable Development in Early Childhood (ERS-SDEC), developed as part of the Education for Sustainable Development World Project led by OMEP, is used to assess sustainable development practices in early childhood education settings (Siraj-Blatchford & Pramling Samuelsson, 2016).In addition, observation-based checklists have been used to examine practices related to sustainable development. For example, Roohi (2018) used an observation checklist covering environmental, social, and economic dimensions to examine the activities of preschool children related to sustainable development. However, this tool was developed for the purpose of collecting data within a specific research context and is not based on a systematic scale development process or validity and reliability analyses.

Upon examining all these assessment tools, it becomes evident that a significant portion of them focuses on children's environmental knowledge, attitudes, or awareness levels, while tools that directly assess behaviors related to sustainability are limited. Furthermore, it is evident that research has largely focused on the environmental dimension of sustainability, whereas the socio-cultural and economic dimensions have been addressed to a relatively limited extent in the literature (Güler Yıldız et al., 2021; Hedefalk et al., 2015). Sustainable development, however, is defined as a holistic approach that addresses environmental, socio-cultural, and economic dimensions together (WCED, 1987; Purvis et al., 2019). Nevertheless, it is noteworthy that most of the measurement tools developed in the literature do not assess these dimensions together within a systematic and holistic framework. This situation highlights the need for assessment tools capable of tracking the development of sustainability practices in early childhood education settings by addressing the environmental, sociocultural, and economic dimensions of sustainability together (Kahriman-Pamuk et al., 2019).

Somerville & Williams (2015) note that while the theoretical approaches developed in the field of sustainability education provide strong conceptual guidance, they fall short in generating concrete and observable indicators that can be used in practice. Similarly, while the literature on sustainability education features robust theoretical debates, it is also noted that these approaches have been limited in their translation into systematic assessment tools in practical contexts (Davis & Elliott, 2023; Sterling, 2001; Wals, 2012). Ardoin&Bowers, (2020) emphasize that while there are tools to measure knowledge, attitudes, and awareness in environmental education, there are fewer tools designed to measure behavioral change. In addition, it is noted that there is a need for assessment tools to evaluate interventions in early childhood sustainability research (Siraj-Blatchford & Pramling Samuelsson, 2016).

Davis & Elliott (2014) state that children make sense of abstract environmental concepts by concretizing them through nature-related experiences-such as those involving air, water, and soil-in their daily lives. Siraj-Blatchford & Pramling Samuelsson (2016) emphasize that learning within the ECEfS framework occurs through play and social interaction. Consequently, this demonstrates that observation-based assessment is appropriate for learning outcomes. Systematic observations and analysis become crucial for identifying children's thoughts and actions related to sustainability in their daily lives (Engdahl, 2015).

The checklist developed for this purpose is intended not only as an assessment tool but also to assist teachers in the planning processes of their sustainability-based practices.

2. Methodology

This study tests the construct validity of a checklist developed for children aged 60-72 months based on the 7R framework. A quantitative research design was used in the study. The developed checklist is grounded in a theoretical foundation. Confirmatory Factor Analysis (CFA) is used to test whether the factor structure of measurement instruments developed based on a theoretical foundation aligns with the data (Brown, 2015; Byrne, 2013; Hair et al., 2019). In this study, CFA was used to establish the construct validity of the checklist.

2.1. Participants

The study sample consists of 448 children aged 60–72 months who were attending preschool institutions during the spring semester of the 2024–2025 academic year. Data were collected from a total of 448 students, of whom 243 were female (54.24%) and 205 were male (45.76%). Within the scope of the study, data were also collected from a total of 61 teachers**.** Each teacher has observed between 5 and 15 children. The number of classes taught by the teachers varies.

Teachers working in preschool institutions across different provinces completed the checklist by observing the children in their classrooms. The checklists were completed by classroom teachers who interacted with the children throughout the academic year. Since the teachers had the opportunity to get to know the children starting in September and to observe them during classroom activities, daily routines, and social interactions, the assessments were based on long-term teacher observations rather than a single observation session. Accordingly, the teachers completed the checklist for each child based on observations made throughout the academic year and submitted the completed forms to the researcher in May.

The items on the checklist consist of clear statements that allow for binary assessment, designed to identify observable indicators of sustainability in children. Consequently, the teachers did not receive extensive training on scoring. However, before the checklists were distributed, the purpose of the tool, the method of assessment, and the requirement to mark each item based on long-term classroom observations of the children were explained to the teachers. Teachers were asked to evaluate each item based on their experience observing the children-from the beginning of the academic year-within the context of classroom activities, daily routines, and social interactions.

Comrey & Lee (1992) recommend a participant-to-item ratio of at least 5:1 for factor analysis during the development of measurement instruments. Additionally, samples exceeding 300 are considered sufficient (Worthington & Whittaker, 2006). In sample selection, a non-probability sampling method was used (Etikan et al., 2016). The researcher's ability to collaborate with teachers they could access during the data collection process influenced the choice of this method. Furthermore, teachers participated in the study on a voluntary basis.

2.2. The Development Process of the Checklist

In this section, the development process of the checklist is explained in a step-by-step manner, beginning with the establishment of the theoretical foundation and continuing through item development and validity and reliability studies.

Determining the Conceptual Framework

The 7R framework holistically addresses the environmental, socio-cultural, and economic dimensions of sustainable development at the early childhood level and defines these dimensions through observable behaviors that children can

demonstrate in their daily lives (Borg & Pramling Samuelsson, 2022; UNESCO, 2017; Siraj-Blatchford & Pramling Samuelsson, 2016). In the development of the checklist, the 7R framework proposed by OMEP served as the theoretical foundation and was used as a reference in structuring the dimensions of the developed checklist.Each dimension within the 7R framework has been translated into concrete behavioral indicators, taking into account the developmental characteristics of children aged 60–72 months and observable behavioral patterns in preschool educational settings. The literature indicates that the 7R framework can be utilized to interpret the environmental, economic, and socio-cultural dimensions of sustainable development in early childhood through the lens of children's thoughts, awareness, and daily life experiences (Kahriman-Öztürk et al., 2012; Borg & Pramling Samuelsson, 2022). Accordingly, the "Respect" dimension is associated with the child refraining from harming nature, living beings, peers, and shared living spaces; the "Rethink" dimension with asking questions about consumption habits, natural assets, the needs of living beings, and environmental processes; the "Reflect" dimension with demonstrating awareness regarding issues such as waste, pollution, product manufacturing processes, diverse living conditions, and social justice; the "Recycle" dimension with sorting waste, recognizing recycling symbols, and properly using recycling bins; the "Use resources equitably" dimension with sharing food, toys, books, and classroom materials; the "Reduce" dimension with avoiding wastefulness in the use of food, water, electricity, and paper; and the "Reuse" dimension with behaviors involving the repurposing or repairing of waste or unused materials. Thus, the abstract principles of the 7R framework have been transformed into behavioral items that teachers can observe in natural classroom settings and evaluate using a Yes/No format.

Creation of the Item Pool

In creating the pool of topics, the literature on sustainability education in early childhood, the 7R framework, environmental awareness in preschool, resource use, sharing, and recycling was examined in a structured manner (Borg and Pramling Samuelsson, 2022; Engdahl, 2015; Siraj-Blatchford & Pramling Samuelsson, 2016; Van Harskamp et al., 2025). As a result of this review, the dimensions through which sustainability education is addressed in early childhood were examined. Consideration was given to how the 7R principles are related to children's daily life experiences and which behaviors can be considered as observable sustainability indicators in preschool education settings.(Borg & Pramling Samuelsson, 2022; Engdahl, 2015; Siraj-Blatchford & Pramling Samuelsson, 2016; Van Harskamp et al., 2025). The item writing process was also informed by studies addressing children’s environmental behaviors, their relationships with nature, and sustainability-related practices in early childhood contexts (Cengizoğlu, 2013; Evans et al., 2007; Güler Yıldız et al., 2021; Kahriman-Pamuk et al., 2019; Roohi, 2018; Yılmaz & Olgan, 2017).

Based on a review of the literature, observable behavioral indicators that are developmentally appropriate and pedagogically meaningful in early childhood settings have been identified.

Kahriman-Öztürk et al. (2012) classified the 7R principles under the three fundamental dimensions of sustainability, associating the "Reduce" and "Reuse" dimensions with the environmental dimension; "Respect," "Rethink," and "Reflect" with the socio-cultural dimension; and "Recycle" and "Redistribute" with the economic dimension. However, it was acknowledged that these dimensions are not sharply demarcated in early childhood education settings and that certain behavioral indicators may relate to more than one dimension of sustainability. Therefore, in the present study, the 7R dimensions were structured during the creation of the item pool based on both classifications found in the relevant literature and observable behavioral patterns of children aged 60–72 months in natural classroom settings.

These items are organized according to the 7R dimensions as follows: Respect (8 items), Rethink (8 items), Reflect (8 items), Recycle (8 items), Redistribute (7 items), Reduce (6 items), and Reuse (6 items).

Expert Review and Content Validity

Expert opinions were sought to ensure the content validity of the developed checklist. The experts consisted of eight academics specializing in environmental education, sustainability, and assessment. The experts were asked to evaluate the items in terms of: a) the extent to which they represent the relevant 7R dimension, b) their appropriateness for the age group, and c) their linguistic clarity.

In the process of developing measurement instruments, expert evaluation is considered a crucial source of validity evidence for assessing the extent to which items represent the construct of interest, the adequacy of the content coverage, and the clarity of the items. Boateng et al. (2018) note that content validity is associated with how adequately items represent the relevant domain, as well as with content appropriateness and technical quality, and that this process can be supported through expert evaluations. Similarly, Artino et al. (2014) state that expert validation during the development of a measurement instrument can be used to assess the clarity and appropriateness of items regarding the construct being measured. In the present study, expert input was obtained through qualitative feedback rather than a quantitative scoring form. It was observed that the feedback from experts largely converged on similar points; consequently, wording adjustments were made to certain items, and items deemed outside the scope were removed from the checklist. Regarding the item "e.g., takes care not to harm surrounding natural habitats and avoids behaviors that could damage these areas (e.g., exhibits eco-friendly behaviors, protects natural habitats)," all experts commented that it would be more appropriate to have a single statement of judgment (either "takes care not to harm" or "avoids behaviors that could cause harm") rather than two. Accordingly, the item has been revised. Another example is: "Does not bring packaged food to school." It has been recommended that this item be removed, as it is under the parents' control rather than the child's initiative. Thus, a qualitative content validity process was conducted to strengthen the checklist's theoretical integrity, suitability for the target age group, and observational usability.

Data Collection Instrument

In this study, the "Sustainable Development Goals Observation Checklist for Children Aged 60-72 Months" was used as the data collection tool. The checklist was designed to assess children's behaviors related to sustainable development through observation in natural learning environments. The literature indicates that checklists are used to determine the presence of behaviors or characteristics and that items can be structured using binary options such as observed/not observed, yes/no, or present/absent (Vercellotti, 2021). Therefore, the items in the checklist for this study were structured to allow for a "Yes/No" indication of whether or not the children exhibited the relevant behaviors. The checklist was completed by two independent raters. An inter-rater agreement analysis was conducted, and the results are summarized below.

- To enhance the reliability of the checklist used in the study, the children were observed by two independent raters, and each item was coded as "yes" (1) or "no" (0). Cohen’s Kappa coefficient was calculated to determine the level of inter-rater consistency.
- The analysis revealed the level of consistency between the two raters to be: Cohen’s Kappa (κ) = 0.92

When evaluated against generally accepted criteria, this value falls into the "almost perfect agreement" category (Landis & Koch, 1977). This result indicates that the items on the observation form were clear and understandable, and that the raters interpreted the same behaviors in a similar manner.

Data Collection Procedure

After the checklist was prepared, approval was obtained from the ethics committee. In addition, since the checklist was to be implemented in schools, research permits were obtained from the Provincial Directorate of National Education. Data were collected through observations conducted by preschool teachers. Since the items on the checklist consist of clear statements that can be evaluated in a binary manner-aimed at identifying observable indicators of sustainability in children-the teachers were not provided with comprehensive rater training. Nevertheless, the researcher briefed the teachers on the purpose of the checklist, the method for evaluating the items, and the criteria to be considered during observation.

Teachers were asked to observe the children within the natural flow of educational activities-including classroom and outdoor activities, play, peer interactions, meals, toiletry routines, and other daily activities. To ensure consistency in assessment, it was specified that for an item to be marked as "Yes/Observed," the child had to demonstrate the relevant behavior during the natural course of events. Conversely, it was clarified that behaviors emerging solely from direct teacher guidance, reminders, or isolated, coincidental incidents should not be considered observed behaviors. Observations were conducted within the children's daily educational routine, with a strong emphasis placed on avoiding any intervention or direction.

The fact that teachers were able to get to know the children from the beginning of the academic year and observe them in various contexts allowed assessments to be based on long-term behavioral patterns rather than a single moment of observation. Basing assessment on behavioral patterns observed across different situations-rather than on a momentary impression-is considered a factor that supports measurement reliability (Bakeman & Gottman, 1997; Mash & Hunsley, 2005). Accordingly, teachers completed the checklist for each child based on observations made throughout the academic year and submitted the completed forms to the researcher in May.

2.3. Data Analysis

Validity and Reliability of the Checklist

Confirmatory factor analysis (CFA) was used in this study. The reason for choosing CFA is not that the 7R is a previously developed measurement tool. The checklist items were developed based on the 7R principles recommended by OMEP for education for sustainable development in early childhood (Kahriman-Öztürk, et.al., 2012). In this study, the 7R principles were treated as an a priori conceptual framework that determined under which behavioral domains the items on the checklist would be organized. Accordingly, observable child behaviors were defined under each principle, and a seven-dimensional measurement structure was established.

For this reason, instead of EFA-which aims to discover the factor structure from the data-CFA was chosen for this study to test the extent to which the predefined seven-dimensional structure is supported by the data. EFA is a method used in situations where there is insufficient theoretical or conceptual grounding and the factor structure cannot be defined in advance as a hypothesis. It has been noted that using EFA in studies where a predefined and theoretically justified structure exists is methodologically inconsistent and analytically inappropriate (Worthington & Whittaker, 2006). In

contrast, CFA allows researchers to test the extent to which the number of factors and item-dimension relationships predefined by the researcher are supported by empirical data (Brown, 2015; Kline, 2016; Hair et al., 2019). Similarly, Büyüköztürk (2017) notes that confirmatory factor analysis is an appropriate method for testing theoretically predefined measurement models.

Each item on the checklist is structured as a binary question to determine whether the relevant behavior was observed in the child. Accordingly, the "Yes/Observed" response indicates that the behavior was observed, while the "No/Not Observed" response indicates that the behavior was not observed. During the scoring process, "Yes/Observed" responses are coded as 1, and "No/Not Observed" responses are coded as 0. Confirmatory factor analysis (CFA) is a widely used method, particularly for measurement instruments where items are scored dichotomously and represent multiple theoretical sub-dimensions, to test the structural validity and fit of the proposed measurement model (Brown, 2015; Byrne, 2013; Hair et al., 2019; Kline, 2016; Wirth & Edwards, 2007). Therefore, in this study, CFA was conducted not to validate the 7R framework itself, but to evaluate the predicted seven-dimensional item structure of the checklist developed based on this frameworkPrevious studies have also used CFA to evaluate the structural validity of theoretically grounded measurement tools. For instance, Alaçam (2022) used CFA to examine the structural validity of a measurement tool developed for portfolio evaluation after the item development process had been shaped by the literature and expert review. In a similar way, Hui et al. (2025) tested the proposed factor structure of the teacher-completed Child Observation Checklist (COC) through CFA and reported evidence supporting the instrument's structural validity.

Upon examining all these examples, it becomes clear that there are no methodological obstacles to applying CFA to establish the construct validity of a theoretically defined measurement instrument (Kline, 2016).

The KR-20 coefficient is an internal consistency coefficient used for dichotomously scored items; it is noted that for items coded on a 0/1 scale, the KR-20 can be interpreted in the same way as Cronbach's alpha (Kuder & Richardson, 1937; Cronbach, 1951). Accordingly, internal consistency coefficients were reported for the general checklist and its subscales. Additionally, considering the multidimensional structure of the checklist based on the 7R framework, a composite reliability coefficient that accounts for different factor loadings was calculated within the scope of CFA. Composite reliability is a measure widely used in reflective measurement models to assess internal consistency reliability (Fornell & Larcker, 1981; Raykov, 1997; Hair et al., 2019).

The analyses were conducted using the Maximum Likelihood (ML) estimation method via the AMOS software on data obtained from 448 participants. Given the binary scoring structure of the items and the possibility of deviation from a normal distribution, the methodological rationale for using the ML estimation method-which is based on the assumption of continuous data-was also evaluated. Although it is known that the ML method relies on the multivariate normality assumption, simulation studies show that, particularly with sufficient sample sizes, ML estimators can be relatively robust to moderate violations of the normality assumption (Finney & DiStefano, 2006; West et al., 1995). In this study, parameter estimates and standard errors were further examined using the bootstrapping method to support the robustness of the findings. The fact that the bootstrapping results were largely consistent with the ML-based outputs was interpreted as additional evidence that deviations from the normality assumption did not lead to significant bias in the model estimates. The use of the bootstrapping approach alongside ML estimation is consistent with practices

recommended in the structural equation modeling literature for assessing the effects of assumption violations (Byrne, 2013; Nevitt & Hancock, 2001).

3. Results

This section presents the CFA results, model fit indices, factor loadings, and validity and reliability analyses for the observation checklist that was developed.

3.1. Findings Related to the Initial Model

The analysis process began with a total of 51 items developed based on the 7R framework of sustainability. Accordingly, the CFA results and goodness-of-fit indices for the initial model are presented in Table 1.

Table 1
Fit Indices for the Initial 51-Item, Seven-Dimension Checklist Based on the 7R Framework

Fit Index	Value	Recommended Threshold	Interpretation
χ^2 (CMIN)	5692.19	-	-
df	1203	-	-
χ^2/df	4.73	< 5.00	Acceptable
CFI	.67	> .95	Low
TLI	.65	> .90	Low
SRMR	.11	< .08	Low
RMSEA	.09	< .06	Low
PClose	.00	> .05	Very low

As reported in Table 1, the chi-square to degrees of freedom ratio for the initial model was 4.73, exceeding the acceptable range. The Comparative Fit Index (CFI = .67) and Tucker–Lewis Index (TLI = .65) also indicated weak model fit. In addition, several items showed low factor loadings (λ < .40), and high error covariances were observed between some item pairs (see Table 2). The RMSEA (.091) and SRMR (.110) values were likewise above the recommended thresholds presented in Table 1 (Hu & Bentler, 1999). Taken together, these results suggest that the initial model did not fit the data at an acceptable level.

To better understand the sources of this poor fit, the standardized factor loadings and item-level error covariances were examined in detail. The standardized factor loadings for the initial model are presented in Table 2.

Table 2
Standardized Factor Loadings for the Initial Seven-Dimension Checklist Structure Based on the 7R Framework

Item	Factor	Standardized Factor Loading
Respect08	F1	,646
Respect 07	F1	,713
Respect06	F1	,690
Respect05	F1	,731
Respect04	F1	,388
Respect03	F1	,551
Respect02	F1	,695
Respect01	F1	,716
Rethink08	F2	,126
Rethink07	F2	,755
Rethink06	F2	,766
Rethink05	F2	,805
Rethink04	F2	,795
Rethink03	F2	,797
Rethink02	F2	,430

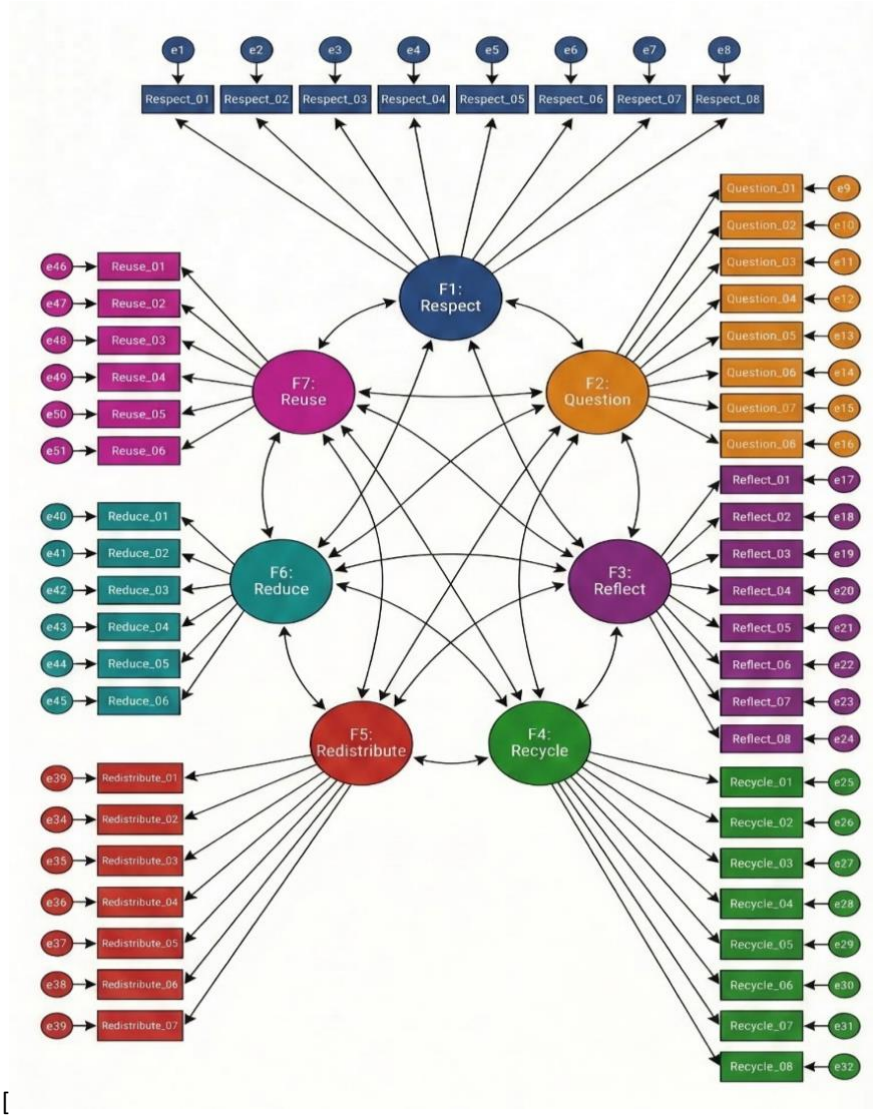
Rethink01	F2	,456
Reflect08	F3	,693
Reflect07	F3	,654
Reflect06	F3	,559
Reflect05	F3	,766
Reflect04	F3	,694
Reflect03	F3	,726
Reflect02	F3	,726
Reflect01	F3	,538
Recycle08	F4	,423
Recycle07	F4	,381
Recycle06	F4	,833
Recycle05	F4	,778
Recycle04	F4	,761
Recycle03	F4	,660
Recycle02	F4	,719
Recycle01	F4	,656
Redistribute01	F5	,372
Redistribute02	F5	,362
Redistribute03	F5	,312
Redistribute04	F5	,398
Redistribute05	F5	,870
Redistribute06	F5	,884
Redistribute07	F5	,749
Reduce01	F6	,415
Reduce02	F6	,549
Reduce03	F6	,677
Reduce04	F6	,731
Reduce05	F6	,764
Reduce06	F6	,500
Reuse01	F7	,669
Reuse02	F7	,708
Reuse03	F7	,761
Reuse04	F7	,722
Reuse05	F7	,809
Reuse06	F7	,612

As presented in Table 2, several items had factor loadings below the recommended threshold of λ = .40. More specifically, Respect04 in the F1 subdimension (λ = .388), Rethink08 in the F2 subdimension (λ = .126), and Recycle07 in the F4 subdimension (λ = .381) fell below this cutoff. A similar pattern was observed in the F5 subdimension, where Redistribute01 (λ = .372), Redistribute02 (λ = .362), and Redistribute03 (λ = .312) also showed loadings below the recommended level. In addition, although it was very close to the threshold, Redistribute04 (λ = .398) was also considered to have a relatively weak factor loading.

The literature suggests that factor loadings below .40 do not provide adequate representation of the relevant latent construct and may weaken overall model fit (Hair et al., 2019). From this perspective, the poor fit indices observed in the initial model appeared to be related, in part, to the weak loadings of several items as well as to the presence of high error covariances. In response to these findings, an iterative revision process was initiated while maintaining the theoretical coherence of the model.

The initially proposed seven-dimension structure and the relationships between the observed items and latent variables are shown in Figure 1. As illustrated in the figure, each item loads onto its corresponding dimension, and the structural relationships among the dimensions are also displayed. In this sense, Figure 1 represents the factor structure of the initial 7R-based model and shows the configuration of the model before the revision process began.

Figure 1
Initial Seven-Factor CFA Model Before Revision



[ ]

The literature suggests that factor loadings below .40 do not adequately represent the relevant latent construct and may weaken overall model fit (Hair et al., 2019). In the initial model, the weak loadings of several items together with high error covariances appeared to contribute to the poor fit indices. To address this issue, an iterative revision process was initiated while preserving the theoretical coherence of the model. Figure 1 presents the initial seven-dimension measurement structure of the checklist, including the standardized loadings of the observed variables on their corresponding latent dimensions and the correlations among those dimensions. When the fit indices reported in Table 1, the factor loadings shown in Table 2, and the overall structure displayed in Figure 1 are considered together, the findings indicate that the initial model required further refinement.

3.2. Model Revision Process and Final Structure

In structural equation modeling studies, a stepwise and iterative item deletion process based on theoretical foundations is recommended to further improve model fit (Hair et al., 2019; Hu & Bentler, 1999). This approach begins with the evaluation of items identified as negatively affecting model fit at each step. The analysis process then continues with the retesting of the model. In this regard, theoretical consistency was maintained in the present study to improve model

fit and strengthen structural validity. Items were not removed en masse; instead, items identified as problematic based on criteria established at each step were evaluated, and the model was retested.

Two criteria guided the item removal process. First, statistical indicators were taken into account, including factor loadings below .40 and high error covariances suggested by modification indices. Second, theoretical consistency was considered by examining, through expert judgment, whether removing a statistically problematic item would reduce the content validity of the checklist.

Despite the item removal process, the originally envisaged 7R-based structure was preserved. Following the revisions, all seven dimensions-Respect, Rethink, Reflect, Recycle, Redistribute, Reduce, and Reuse-remained in the final model, and each dimension was represented by at least three items. This demonstrates that the item removal process was not intended to alter the theoretical structure, but rather to ensure that each dimension was represented by stronger and more consistent indicators. Furthermore, expert opinions were sought to assess whether the removed items narrowed the scope of the relevant dimension; thus, while improving model fit, care was taken to preserve the 7R-based conceptual integrity of the checklist.

As a result of this evaluation process, 16 items were removed from the model, and the seven-dimension structure was retained in a final form consisting of 35 items. Throughout the revision process, care was taken to ensure that the core indicators of each factor remained represented and that the overall content coverage of the checklist was preserved. The relatively strong correlations observed among the dimensions suggested that they might be accounted for under a common higher-order structure. The literature supports the use of second-order factor models when the shared variance among multiple first-order factors can be theoretically explained by a broader latent construct (Brown, 2015; Byrne, 2013; Hair et al., 2019; Kline, 2016).

3.3. Goodness-of-Fit Results of the Final Model

The goodness-of-fit indices obtained from the second-order CFA conducted for the 35-item checklist indicated that the model fits the data at an acceptable level. The χ^2/df ratio of 3.03 falls within the range (3 to 5) specified by Marsh & Hocevar (1985) as acceptable. The result obtained in this study was found to be 0.067. However, the fact that the SRMR value (0.04) is below 0.08 indicates a good fit (Browne & Cudeck, 1993; Hu & Bentler, 1999).

Hu & Bentler (1999) note in their study on structural equation modeling that CFI and TLI values of .90 or higher are recommended as an acceptable level of model fit. The CFI (0.89) and TLI (0.87) values obtained in this study were found to be quite close to the acceptable fit threshold. It is therefore evident that the model's overall fit can be considered within acceptable limits. Based on this, we can conclude that the checklist developed supports the theoretically predicted second-order structure statistically. Detailed fit indices for the final model are presented in Table 3.

Table 3
Goodness-of-Fit Results of the Second-Order Confirmatory Factor Analysis

Fit Index	Obtained Value	Threshold Value (Acceptable)	Interpretation
χ^2/df	1573.42 / 519 = 3.03	< 5.00	Acceptable
RMSEA	0.07	< 0.08	Good fit
CFI	0.89	≈ 0.90	Close to acceptable threshold
TLI	0.87	≈ 0.90	Close to acceptable threshold
SRMR	0.04	< 0.08	Good fit

As reported in Table 3, the chi-square ratio remained below 5, and both the RMSEA and SRMR values met the recommended cutoff criteria, indicating an acceptable fit between the model and the data. Although the CFI and TLI values were slightly below the .90 threshold, the literature suggests that, in complex models involving a large number of items and dichotomously scored data, values of this kind may still be regarded as close to acceptable limits (Marsh et al., 2004).

3.4. Factor Loadings and Structural Relationships

In the final model, the standardized factor loadings ranged from .46 to .90 (see Figure 1). As presented in Table 4, all item loadings (*t* values) were statistically significant ($p < .001$). In general, factor loadings above .40 suggest that the items have adequate discriminative capacity in representing their corresponding construct (Hair et al., 2019).

Table 4

Standardized Factor Loadings of the Final Model with 7 Dimensions and 35 Items

Factor	Item	Standardized Factor Loading
Sustainable Development Behaviors (F8)	Respect (F1)	,399
	Rethink (F2)	,936
	Reflect (F3)	,937
	Recycle (F4)	,789
	Redistribute (F5)	,428
	Reduce (F6)	,848
	Reuse (F7)	,700
Respect (F1)	Respect07	,783
	Respect06	,566
	Respect05	,647
	Respect02	,717
	Respect01	,741
Rethink (F2)	Rethink07	,727
	Rethink06	,762
	Rethink05	,761
	Rethink04	,774
	Rethink03	,802
	Rethink01	,455
Reflect (F3)	Reflect08	,612
	Reflect07	,643
	Reflect05	,764
	Reflect04	,590
	Reflect03	,719
	Reflect02	,787
Recycle (F4)	Recycle06	,810
	Recycle05	,746
	Recycle04	,765
	Recycle03	,662
	Recycle02	,775
	Recycle01	,672
Redistribute (F5)	Redistribute05	,895
	Redistribute06	,893
	Redistribute07	,755
Reduce (F6)	Reduce03	,633
	Reduce04	,616
	Reduce05	,704
Reuse (F7)	Reuse01	,603
	Reuse02	,627
	Reuse03	,675
	Reuse04	,772

Factor	Item	Standardized Factor Loading
	Reuse05	,837
	Reuse06	,733

According to the second-order CFA results, the higher-order factor conceptualized as behaviors aligned with sustainable development showed strong relationships with the first-order dimensions. The standardized regression coefficients linking the higher-order factor to the subdimensions ranged from .40 to .94. Among these, Reflect (β = .94) and Rethink (β = .94) emerged as the dimensions most strongly associated with the overarching sustainable development construct (see Table 4).

3.5. Validity and Reliability Results

The construct validity and reliability of the scale were evaluated using Composite Reliability (CR), Average Variance Extracted (AVE), and discriminant validity indicators, including MSV and HTMT (see Table 5).

Table 5
Validity and Reliability Analysis Result

Factor / Construct	CR	AVE	Discriminant Validity
Behaviors Aligned with Sustainable Development (Overall)	0.89	0.56	Established

Note. The recommended cutoff values were .70 for Composite Reliability and .50 for Average Variance Extracted (Fornell & Larcker, 1981).

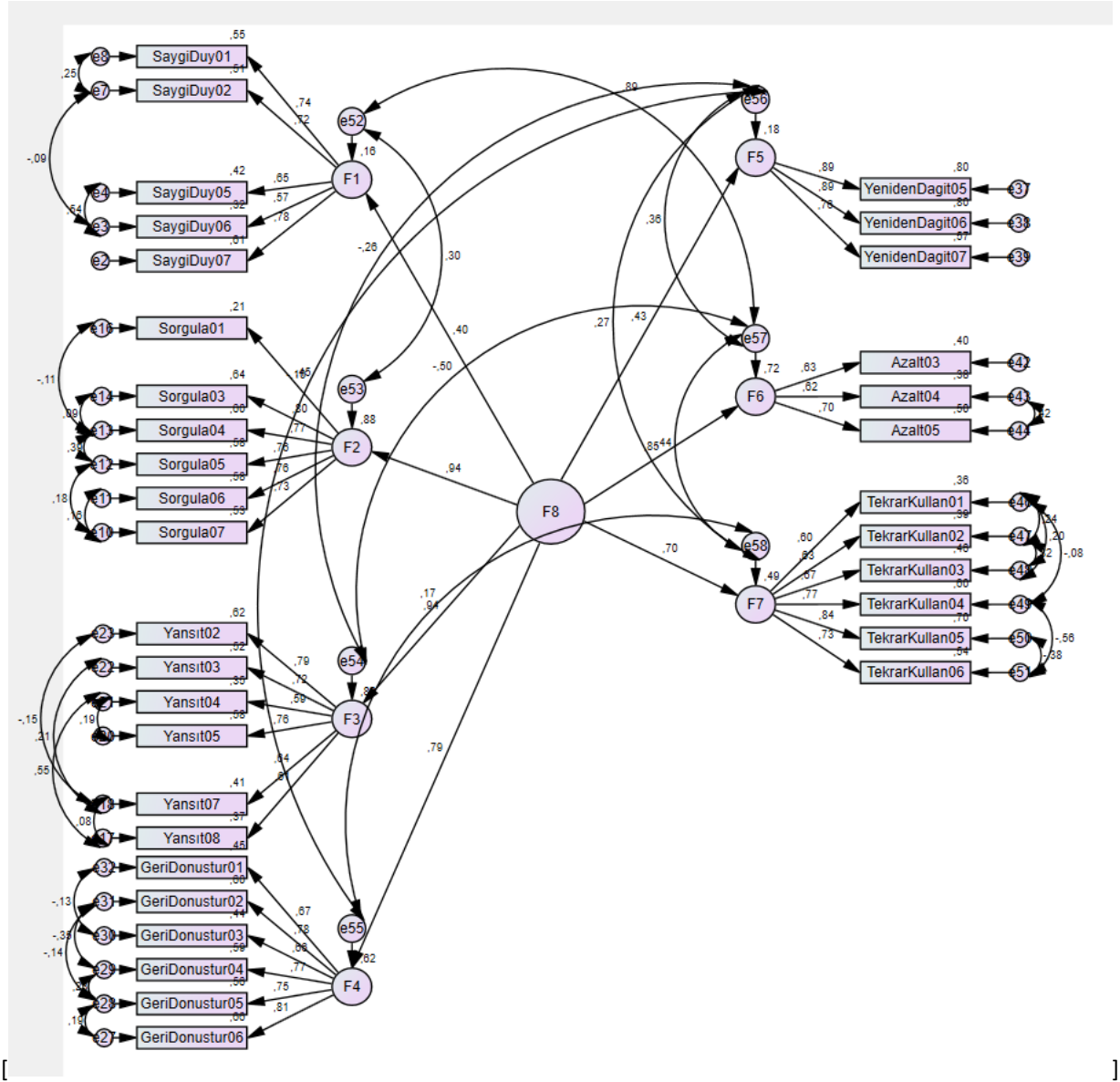
Convergent Validity: The CR value calculated for the general factor (.892) was well above the .70 threshold. Similarly, the AVE value (.561) exceeded the recommended cutoff of .50 (Fornell & Larcker, 1981). Together, these findings indicate that the scale demonstrated adequate convergent validity and internal consistency.

Discriminant Validity: The results of the analysis showed that the Maximum Shared Variance (MSV) value (.000) was lower than the Average Variance Explained (AVE) value (MSV < AVE). The MSV value of .000 is due to the second-level hierarchical structure of the final model. In the model, all first-level factors were defined under a single top-level factor, and no free covariances were estimated among the first-level factors. Since MSV represents the highest squared correlation among latent structures, the absence of direct free covariances between factors in such hierarchical models can lead to the MSV value being reported as .000 (Hair et al., 2019; Marsh et al., 2014). Therefore, this value was interpreted as a result of the model's structural specification rather than a computational error. Furthermore, the Heterotrait-Monotrait (HTMT) ratio analysis revealed that the inter-factor values remained below the 0.85/0.90 thresholds, and no validity warning was present in the analysis report (Henseler et al., 2015). Consequently, the assessment of discriminant validity was not limited to the MSV value but was interpreted in conjunction with the HTMT results.

In summary, the final model, reduced from 51 to 35 items, demonstrated a valid and reliable structure in terms of both statistical fit and theoretical coherence. Following the goodness-of-fit and construct validity findings, Figure 2 displays the 35-item second-order confirmatory factor analysis model. As shown in the figure, the first-order dimensions are hierarchically organized under the higher-order factor conceptualized as behaviors aligned with sustainable development, and the standardized item loadings across the factors show an acceptable pattern. In this sense, Figure 2 offers an overall representation of the structural coherence and hierarchical organization of the final model that emerged through the revision process.

Figure 2

Final Second-Order CFA Model After Revision



As illustrated in Figure 2, the seven dimensions are organized consistently under the higher-order factor conceptualized as behaviors aligned with sustainable development, and the standardized item loadings on their respective dimensions follow an acceptable pattern. After the revision process, the model reached a more parsimonious structure, with the relationships among the first-order factors accounted for by the higher-order factor. Taken together, this structure supports both the hierarchical coherence of the scale and the statistical consistency of the measurement model.

3.6. Internal Consistency Reliability Results

Table 6 reports the Cronbach's α coefficients for each factor as well as for the overall scale, providing evidence on the internal consistency reliability of the instrument.

Table 6
Cronbach's α Values of the Overall Checklist and Its Subfactors

Factor	Number of Items	Cronbach α
Respect	5	0.83
Rethink	6	0.87
Reflect	6	0.86
Recycle	6	0.87
Redistribute	3	0.88
Reduce	3	0.76
Reuse	6	0.86
Overall Checklist	35	0.94

The literature generally suggests that a Cronbach's α value of at least .70 indicates acceptable internal consistency reliability (Nunnally & Bernstein, 1994). It was determined that the Cronbach's alpha coefficients for the subscales ranged from .76 to .88. These values indicate that the internal consistency of the subscales is adequate. The Cronbach's alpha coefficient for the checklist as a whole was calculated as .94. This result indicates a high level of internal consistency across the entire checklist.

4. Discussion

In this study, an observation-based checklist was developed based on the 7R framework proposed by OMEP to assess behaviors aligned with sustainable development in children aged 60–72 months. The findings indicate that the instrument, which initially consisted of 51 items, was refined through a series of revisions to reach its final form of 35 items, while maintaining the conceptually anticipated seven-dimensional structure. In its final form, all seven dimensions-Respect, Rethink, Reflect, Recycle, Redistribute, Reduce, and Reuse-were included, with each dimension represented by at least three items. This result demonstrates that the 7R framework provides a structure that can be used to assess sustainability behaviors in early childhood through observation. This finding is consistent with studies indicating that the 7R framework can holistically address the environmental, sociocultural, and economic dimensions of sustainable development at the early childhood level (Borg & Pramling Samuelsson, 2022; Kahriman-Öztürk et. al., 2012; Siraj-Blatchford & Pramling Samuelsson, 2016).

It is emphasized that statistical fit must be evaluated in conjunction with the theoretical framework in the development of measurement instruments (Brown, 2015; Byrne, 2013; Kline, 2016). Accordingly, the results of confirmatory factor analysis revealed that the initial model did not fit the data adequately; in contrast, the final model achieved an acceptable level of fit following revisions. This finding indicates that the initial version of the theoretically constructed structure was not fully supported by the data; however, through the gradual removal of items exhibiting low factor loadings and high error correlations, a simpler and psychometrically more consistent structure was achieved. In the item removal process, consideration was given not only to statistical indicators but also to the items' ability to represent the relevant 7R dimension and to expert opinions, which contributed to preserving the content validity of the checklist. From this perspective, the revision process conducted in this study can be regarded as a procedure aimed at preserving the checklist's 7R foundation while improving model fit.

One of the study's key findings is the integration of the seven subdimensions under a more general overarching structure. Second-order CFA results showed that these subdimensions are grouped under a more general structure that

could be termed 7R, or "behaviors consistent with sustainable development." This finding is consistent with explanations suggesting that the environmental, sociocultural, and economic dimensions of sustainable development are not entirely independent of one another but rather form an interconnected and holistic structure (Kioupi & Voulvoulis, 2019; Purvis et al., 2019). This finding also supports the view that, in early childhood, the 7R framework can be approached not only as a pedagogical orientation but also as a multidimensional structure that can be assessed through observable child behaviors (Kahriman-Öztürk et al., 2012; Özkan et al., 2020).

The reliability findings indicate that the final checklist has an acceptable level of internal consistency. It is noted that internal consistency coefficients of .70 or higher are generally considered acceptable for measurement instruments (Nunnally & Bernstein, 1994). The fact that the internal consistency coefficients for the subscales are at an adequate level supports the notion that each R dimension consists of items that are internally consistent. The high internal consistency coefficient for the overall checklist, meanwhile, indicates that the 35-item form provides a consistent structure for comprehensively assessing children's behaviors aligned with sustainable development.

The checklist developed in this study differs in certain respects from existing assessment tools in the literature. A significant portion of the existing tools related to sustainability and environmental education in early childhood focus on assessing children's environmental knowledge, attitudes, awareness, or perceptions (Evans et al., 2007; Gökçeli & Kandır, 2016; Leeming et al., 1995; Musser & Diamond, 1999; Özkan et al., 2020). This study, however, aims to assess children's sustainability-related behaviors in natural classroom settings based on teacher observations. Furthermore, structuring the checklist based on the 7R framework ensures that the tool addresses not only environmental behaviors but also behaviors related to the sociocultural and economic dimensions of sustainable development. In this regard, the tool contributes to a more holistic and observation-based assessment of sustainability behaviors in early childhood, as it enables children's sustainability-related behaviors to be directly observed across environmental, social, and economic dimensions and supports the collection of reliable child-level data.

Checklists are tools used to determine the presence of a behavior or trait, and their items can be organized in a binary format such as "observed/not observed," "yes/no," or "present/absent" (Vercellotti, 2021). The structure of the checklist as an observation-based tool scored on a binary scale ("Yes-Observed"/"No-not Observed") is considered consistent with the study's purpose. In this study as well, for each item, the "Yes/Observed" option is selected if the child exhibits the relevant behavior in a natural classroom setting, and the "No/Not Observed" option is selected if the child does not exhibit it. Subscale scores can be calculated by summing the behaviors observed within the relevant dimension. The total score can be interpreted as a general observation score indicating the extent to which the child demonstrates indicators of behavior aligned with sustainable development across the 35 items. Higher scores indicate that the child exhibits more behavioral indicators aligned with sustainable development. However, these scores should not be interpreted as diagnostic or classificatory results, but rather as a pedagogical indicator to help teachers monitor children's sustainability behaviors and plan their educational processes.

While the literature frequently emphasizes the importance of sustainability education in early childhood, it also notes that there are limited tools for assessing children's behaviors in this area within natural educational settings (Ardoin&Bowers, 2020; Davis & Elliott, 2023; Siraj-Blatchford & Pramling Samuelsson, 2016; Van Harskamp et al., 2025). The primary contribution of this study to the field is that it presents a checklist based on the 7R framework, which enables the assessment of sustainability behaviors by translating them into observable behavioral indicators at the early

childhood level. In this context, the developed checklist can help teachers monitor children's sustainability behaviors within daily classroom life and identify which "R" dimensions require additional support. Therefore, the developed checklist can be considered not only a research-oriented measurement tool but also a practical assessment tool that can contribute to the planning of sustainability-based activities in early childhood education settings.

Limitations

This study has certain limitations. Although the sample size was sufficient for confirmatory factor analysis, the study group was selected using a convenience sampling method; consequently, the generalizability of the findings is limited in this regard. A second limitation is that the study did not examine whether scores obtained from the developed checklist differed based on variables such as the children's gender, the teachers' years of service, or the type of school. Since the primary aim of this research was to develop an observation-based checklist to assess behaviors related to sustainable development in children aged 60-72 months-and to provide initial evidence regarding the tool's validity and reliability-group comparisons based on these variables were excluded from the scope of the study. Future studies are encouraged to examine the checklist scores in relation to variables such as child gender, teacher experience, and school type. Another limitation stems from the fact that the teachers were already familiar with the children over an extended period. While this familiarity supports the assessment process, future studies involving different observers to examine inter-rater reliability could provide stronger evidence regarding the tool's validity.

5. Conclusions and Recommendations

In this study, the theoretical structure of an observation-based checklist developed based on the 7R framework proposed by OMEP was examined through confirmatory factor analysis (CFA). The results showed that sustainable development-related behaviors were meaningfully organized within a seven-dimensional structure and that these dimensions were integrated under a second-order construct. In addition, the model fit indices (χ^2/df = 3.03, RMSEA = .067, SRMR = .04) indicated that the proposed model achieved an acceptable level of fit. Taken together, these findings suggest that the developed instrument is consistent with its theoretically defined structure.

The final version of the checklist has been reduced from 51 items to 35, while maintaining its theoretical integrity throughout the process. This has resulted in a shorter and more practical instrument. This reduction in the number of items has made it easier for teachers to use the checklist during observations as part of their daily teaching routine.

The findings regarding validity and reliability indicate that the developed checklist is capable of consistently assessing behaviors related to sustainable development. Observation-based tools play a crucial role in monitoring children's behaviors in their daily routines during early childhood. The checklist developed in this study was designed to enable teachers observe and track behaviors related to sustainability and plan the educational process accordingly. Future studies could be conducted in different cultural contexts. Additionally, data collection involving multiple observers and longitudinal designs could provide stronger evidence regarding the instrument's validity.

Declaration of Conflicts of Interest

The authors declare that they have no conflicts of interest.

Declaration of Generative AI Use

During the writing process of this article, ChatGPT was used solely for grammar checking.

Ethical Statement

This study involved collecting data from human participants through a questionnaire. Ethical approval for the study was obtained from the Scientific Research and Publication Ethics Committee of Fatih Sultan Mehmet Foundation University, with the decision number 36/19 and dated 06.06.2024.

Author Contributions

Prof. Dr. Fatma Alisinanoğlu and Assoc. Prof. Dr. Simge Yılmaz Uysal contributed to developing the theoretical framework and designing the study. Gülcan Gül Toker was responsible for planning and conducting the data collection and analysis processes. Prof. Dr. Fatma Alisinanoğlu and Assoc. Prof. Dr. Simge Yılmaz Uysal contributed to the discussion section. The first draft of the manuscript was written by Gülcan Gül Toker, and all authors reviewed each stage of the study. All authors read and approved the final version of the manuscript.

Acknowledgements

This study was produced from the data of the doctoral dissertation titled “The Effect of Story-Based Activities Aligned with the Sustainable Development Goals on the Self-Regulation Skills of Children Aged 60–72 Months,” conducted at the Graduate Education Institute of Fatih Sultan Mehmet Foundation University**.**

References

Alaçam, N. (2022). *A comparative investigation of early childhood teachers’ portfolio practices: Cases from Turkey and the United States* [Doctoral dissertation, Middle East Technical University]. YÖK Ulusal Tez Merkezi. https://tez.yok.gov.tr/UlusalTezMerkezi/tezDetay.jsp?id=721479

Artino, A. R., Jr., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher, 36*(6), 463–474. https://doi.org/10.3109/0142159X.2014.889814

Ardoin, N. M., & Bowers, A. W. (2020). Early childhood environmental education: A systematic review of the research literature. *Educational Research Review, 31*, 100353.

Ärlemalm-Hagsér, E. (2012). Lärande för hållbar utveckling i förskolan: Kunskapsinnehåll, delaktighet och aktörskap kommunicerat i text. *Nordisk barnehageforskning*, *5*. https://doi.org/10.7577/nbf.4175.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis* (2nd ed.). Cambridge University Press.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Bhagwanji, Y. (2011). *Early childhood environmental education rating scale*. North American Association for Environmental Education.

Borg, F., & Pramling Samuelsson, I. (2022). Preschool children’s agency in education for sustainability: The case of Sweden. *European Early Childhood Education Research Journal, 30*(1), 147–163. https://doi.org/10.1080/1350293X.2021.2008922

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.

Büyüköztürk, Ş. (2017). *Sosyal bilimler için veri analizi el kitabı* (23. baskı). Pegem Akademi.

Cengizoğlu, S. (2013). *Investigating the potential of an education for sustainable development program on preschool children's perceptions of the human–environment relationship* (Master's thesis, Middle East Technical University). YÖK Ulusal Tez Merkezi. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Chapman, S. N., & O'Gorman, L. (2022). *Transforming learning environments in early childhood contexts through the arts: Responding to the United Nations Sustainable Development Goals*. International Journal of Early Childhood, 54, 37–55. https://doi.org/10.1007/s13158-022-00320-3

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

Cutter-Mackenzie, A., & Edwards, S. (2013). Toward a model for early childhood environmental education: Foregrounding, developing, and connecting knowledge through play-based learning. *The Journal of Environmental Education*, *44*(3), 195-213. https://doi.org/10.1080/00958964.2012.751892

Davis, J. (2009). Revealing the research 'hole' of early childhood education for sustainability: a preliminary survey of the literature. *Environmental Education Research*, *15*(2), 227–241. https://doi.org/10.1080/13504620802710607

Davis, J., & Elliott, S. (Eds.). (2023). *Young children and the environment: Early education for sustainability*. Cambridge University Press.

Davis, J. M., & Elliott, S. (2014). An orientation to early childhood education for sustainability and research – Framing the text. In J. M. Davis & S. Elliott (Eds.), *Research in early childhood education for sustainability: International perspectives and provocations* (pp. 1–16). Routledge. https://doi.org/10.4324/9781315767499

Engdahl, I. (2015). Early childhood education for sustainability: The OMEP world project. *International Journal of Early Childhood, 47*(3), 347–366. https://doi.org/10.1007/s13158-015-0149-6

Engdahl, I., & Rabušicová, M. (2011). Children's voices about the state of the earth. *International Journal of Early Childhood, 43*(2), 153–176. https://doi.org/10.1007/s13158-011-0031-0

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Evans, G. W., Brauchle, G., Haq, A., Stecker, R., Wong, K., & Shapiro, E. (2007). Young children's environmental attitudes and behaviors. *Environment and Behavior, 39*(5), 635–658. https://doi.org/10.1177/0013916506294252

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Information Age Publishing.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Güler Yıldız, T., Öztürk, N., İlhan İyi, T., Aşkar, N., Banko Bal, Ç., Karabekmez, S., & Höl, Ş. (2021). Education for sustainability in early childhood education: A systematic review. *Environmental Education Research, 27*(6), 796–820. https://doi.org/10.1080/13504622.2021.1896680

Gökçeli, F., & Kandır, A. (2016). 48–66 aylık çocuklar için çevresel farkındalık değerlendirme ölçeğinin geliştirilmesi. *Eğitim ve Bilim, 41*(186), 113–126. https://doi.org/10.15390/EB.2016.6393

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

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.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hui, A. N. N., Siu, A. F. Y., Leung, C., & Karnilowicz, W. (2025). Development and validation of Child Observation Checklist for early screening of young children with special needs. *Healthcare, 13*(2), 148. https://doi.org/10.3390/healthcare13020148

Kahriman-Öztürk, D., Olgan, R., & Güler, T. (2012). Preschool Children's Ideas on Sustainable Development: How Preschool Children Perceive Three Pillars of Sustainability with the Regard to 7R. *Educational Sciences: Theory & Practice*, *12*(4).

Kahriman-Pamuk, D., Uzun, N. B., Yıldız, T. G., & Haktanır, G. (2019). Reliability of the Environmental Rating Scale for Sustainable Development in Early Childhood (ERS-SDEC) indicators. *International Journal of Early Childhood, 51*(2), 193–206. https://doi.org/10.1007/s13158-019-00243-6

Kahriman-Pamuk, D., & Pramling Samuelsson, I. (2024). *Nurturing sustainability in toddlerhood: Investigating early childhood teachers' views and daily practices in a Swedish preschool*. Preprints. https://doi.org/10.20944/preprints202410.1744.v1

Kioupi, V., & Voulvoulis, N. (2019). Education for sustainable development: A systemic framework. *International Journal of Sustainability in Higher Education, 20*(4), 744–765. https://doi.org/10.1108/IJSHE-10-2018-0202

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Köklü-Yaylacı, H. & Feriver, Ş. (2020). Erken çocukluk döneminde çevre eğitimi ve sürdürülebilir kalkınma için eğitim. In R. Olgan (Ed)., *Erken çocukluk döneminde çevre eğitimi* (pp.18-40). Pegem Akademi.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151–160. https://doi.org/10.1007/BF02288391

Leeming, F. C., Dwyer, W. O., & Bracken, B. A. (1995). Children's environmental attitude and knowledge scale: Construction and validation. *The journal of environmental education*, *26*(3), 22-31.

Landis, JR ve Koch, GG (1977). Birden fazla gözlemci arasında çoğunluk anlaşmasının değerlendirilmesinde hiyerarşik kappa tipi istatistiklerin bir uygulaması. *Biometrics* , 363-374.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin, 97*(3), 562–582.

Mash, E. J., & Hunsley, J. (2005). Developing guidelines for the evidence-based assessment. *Psychological Assessment, 17*(3), 362–379.

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Musser, L. M., & Diamond, K. E. (1999). The children's attitudes toward the environment scale for preschool children. *The Journal of Environmental Education*, *30*(2), 23-30.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling, 8*(3), 353-377. https://doi.org/10.1207/S15328007SEM0803_2

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Özkan, B., Tuğluk, M. N., & Yiğitalp, N. (2020). Environmental sustainability scale for children 60–72 months old: A validity and reliability study. *Journal of Education and Training Studies, 8*(1), 32–41. doi:10.11114/jets.v8i1.4485

Pramling Samuelsson, I., & Kaga, Y. (Eds.). (2008). *The contribution of early childhood education to a sustainable society*. UNESCO.

Purvis, B., Mao, Y., & Robinson, D. (2019). Three pillars of sustainability. *Sustainability Science, 14*(3), 681–695. https://doi.org/10.1007/s11625-018-0627-5

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173–184. https://doi.org/10.1177/01466216970212006

Roohi, T. (2018). Influence of sustainable development values on young children. *People: International Journal of Social Sciences, 4*(3), 654–670. https://doi.org/10.20319/pijss.2018.43.654670

Samuelsson, I. P., & Park, E. (2017). How to educate children for sustainable learning and for a sustainable world. *International Journal of Early Childhood, 49*(3), 273–285. https://doi.org/10.1007/s13158-017-0197-1

Siraj-Blatchford, J., & Pramling-Samuelsson, I. (2016). Education for sustainable development in early childhood care and education: An introduction. In *International research on education for sustainable development in early childhood* (pp. 1-15). Cham: Springer International Publishing.

Somerville, M., & Williams, C. (2015). Sustainability education in early childhood. *Contemporary Issues in Early Childhood, 16*(2), 102–117. https://doi.org/10.1177/1463949115585658

Sterling, S. (2001). *Sustainable education: Re-visioning learning and change.* Green Books.

UNESCO. (2017). Education for Sustainable Development Goals Learning Objectives. Paris: UNESCO. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000247444 (Accessed: 1 March 2026).

Yılmaz, S., ve Olgan, R. (2017). Okul öncesi dönem çocuklarının doğaya yakınlık (biyofili) seviyelerinin araştırılması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, *13*(3), 1106-1129. https://doi.org/10.17860/mersinefd.328340

Van Harskamp, M., De Maeyer, S., Sass, W., Van Petegem, P., & Boeve-de Pauw, J. (2025). Measurement invariance of the action competence questionnaire. *Environmental Education Research, 31*(3), 498–511. https://doi.org/10.1080/13504622.2024.2382967

Vercellotti, M. L. (2021). Beyond the rubric: Classroom assessment tools and assessment practice. *TESL-EJ, 25*(3), 1–15.

Wals, A. E. (2012). Shaping the education of tomorrow: 2012 full-length report on the UN decade of education for sustainable development.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Sage.

Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: current approaches and future directions. *Psychological Methods*, *12*(1), 58.

World Commission on Environment and Development. (1987). *Our common future*. Oxford University Press.

World Organisation for Early Childhood Education (OMEP). (2017). *OMEP ESD rating scale for sustainable development in early childhood (ERS-SDEC)*. OMEP & UNESCO.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806–838. https://doi.org/10.1177/0011000006288127