



Design Recommendations for a Child-Oriented Christian Education App Based on Usability Evidence

Henoch Juli Christanto ^{1*}

^{1*} Department of Informatics, Faculty of Engineering, Christian University of Indonesia, East Jakarta City, Special Capital Region of Jakarta, Indonesia.

*Corresponding author: henoch.christanto@uki.ac.id

Received: February 12, 2026; Accepted: March 25, 2026; Published: April 1, 2026.

Abstract: Technology has become a routine part of children's learning, including the transmission of religious values through mobile applications. This study evaluates the usability of the Superbook Kids Bible App — a child-oriented Christian learning application — using Nielsen's five usability dimensions: learnability, efficiency, memorability, errors, and satisfaction. A descriptive quantitative survey was conducted with 100 respondents selected through purposive sampling, with sample size determined by Slovin's formula at a 10 percent margin of error. A 25-item questionnaire was validated using Pearson item-total correlations, yielding validity coefficients ranging from 0.598 to 0.824, while reliability testing confirmed good internal consistency. Descriptive results indicate moderate overall usability, with mean scores of 22.40 for learnability, 8.12 for efficiency, 8.65 for memorability, 17.76 for errors, and 14.91 for satisfaction. Category analysis identifies efficiency as the primary bottleneck, with 46 percent of respondents in the low category, followed by learnability at 33 percent, while errors were predominantly moderate at 60 percent. Based on these findings, the study proposes evidence-based design priorities: performance improvements to reduce loading and response delays, streamlined navigation to cut unnecessary steps, and guided onboarding with child-appropriate prompts to support early-stage learnability, alongside clearer feedback and error recovery mechanisms. Future research should combine task-based observation with usage analytics, test interventions such as gamification and personalization, and apply longitudinal designs to examine sustained engagement and learning outcomes in faith-based mobile learning.

Keywords: Mobile Learning Application; Children-Computer Interaction; Child-Oriented Usability; Usability; Religious Education Technology.

1. Introduction

The spread of mobile technology into children's daily routines has made digital media a practical vehicle for education, including the transmission of religious values (El-Asam *et al.*, 2021; Myers & Arterberry, 2022). Children today grow up interacting with touchscreens before they can read, and their learning preferences reflect that reality — they respond more readily to animated, interactive content than to static text or conventional classroom materials (Mertala *et al.*, 2024; Thompson, 2013). Mobile applications have therefore become a viable medium for faith-based instruction, offering content that can be accessed at home, at any hour, and at a pace suited to individual children (Casidy *et al.*, 2016).

The Superbook Kids Bible App is one such application. Developed to support Christian learning for children, it offers animated Bible stories, interactive games, quiz-based activities, and other features designed

to make scripture accessible and memorable. Yet rich multimedia alone does not guarantee that an application works for its intended users (Hiuredhy *et al.*, 2024; Li *et al.*, 2025). Children may struggle with navigation, misread instructions, miss feedback cues, or lose interest mid-session — not because the content is poor, but because the interface creates friction that the content cannot overcome (Christanto *et al.*, 2024a). The scale of the application's adoption makes this concern worth examining carefully. The Google Play listing records more than 10 million downloads, a 4.6-star rating, and approximately 294,000 user reviews (Google Play, 2025). In Indonesia specifically, national data from March 2025 show that 85.78 percent of students aged 5 to 23 years used mobile phones and 83.80 percent used the internet (Ridwan, 2025). An application reaching children at this scale carries a proportional responsibility to function well — not merely to attract downloads, but to sustain learning.

Evaluating usability for child-oriented applications is not equivalent to evaluating usability for adults (El-Asam *et al.*, 2021; Pal & Vanijja, 2020). Children's cognitive development, motor control, and attention patterns shape how they read visual cues, interpret system responses, and carry out tasks such as selecting menus or completing interactive exercises (Hagen *et al.*, 2023). Usability in this context is not a secondary concern — it determines whether the application can deliver on its educational purpose at all (Antonelli *et al.*, 2023; Christanto *et al.*, 2023). Design decisions about button size, text readability, color contrast, interaction pacing, audio clarity, and in-app guidance all affect whether children can stay engaged and learn (Vasalou & Gauthier, 2023).

Despite growing adoption of mobile learning tools, systematic usability evaluations of religious education applications for children remain scarce. Most existing studies focus on general educational apps, adult populations, or broad satisfaction measures, without examining how specific design choices shape children's task performance, error patterns, or faith-based learning engagement (Itzhak *et al.*, 2023; Mavish *et al.*, 2024). Many evaluations also rely exclusively on questionnaires, which carry known limitations when used with children whose self-reporting abilities vary by age and reading level — pointing to the need for child-appropriate evaluation methods that combine subjective feedback with direct behavioral observation (Kakavand *et al.*, 2023; Christanto *et al.*, 2024b).

This study addresses that gap by evaluating children's interactions with the Superbook Kids Bible App and examining how its design shapes usability and learning experience. The specific aims are: (1) to assess the application's usability using established child-centered usability criteria, and (2) to capture user perceptions through structured surveys completed with parental assistance. Results are used to identify design strengths and areas requiring attention across Nielsen's five usability dimensions: learnability, efficiency, memorability, errors, and satisfaction. The practical contribution is a set of concrete improvement priorities for developers of religious educational applications. The academic contribution is empirical evidence on usability evaluation in faith-based learning contexts — a category that remains underrepresented in the mobile learning literature. A well-designed religious learning application does more than perform reliably. When children can navigate it without confusion, complete activities without interruption, and return to it without relearning the interface, the application becomes a genuine tool for values education — one that parents and educators can rely on as a supplement to formal instruction.

2. Related Work

Research on usability evaluation in educational technology has grown substantially over the past decade, yet studies specifically targeting child-oriented religious learning applications remain limited. The existing body of work can be organized around three intersecting themes: usability evaluation frameworks and methods, child-computer interaction in learning contexts, and faith-based educational application design. Nielsen's five usability dimensions — learnability, efficiency, memorability, errors, and satisfaction — remain the most widely applied framework for assessing the quality of interactive systems (Nielsen, 1993, 2005). These dimensions provide a structured basis for diagnosing interaction problems and have been applied across a range of educational platforms. Christanto *et al.* (2023) applied this framework to compare usability in Google Classroom and Microsoft Teams, finding that efficiency and learnability were the dimensions most sensitive to interface differences between platforms. Similarly, Christanto & Sedyono (2020) used Nielsen's dimensions to evaluate an online train ticket reservation system, demonstrating that the interval-based category classification approach produces diagnostic results that can directly inform design priorities. Pal & Vanijja (2020) extended usability evaluation to online learning platforms during the COVID-19 period, using the System Usability Scale alongside the Technology Acceptance Model, and found that perceived usability was strongly linked to continued engagement — a finding directly relevant to child-oriented applications where sustained use is an educational goal. Mavish *et al.* (2024) further demonstrated that usability evaluation combined with design thinking methods can produce actionable redesign recommendations, reinforcing the value of structured assessment as a precursor to interface improvement.

Evaluating usability for children requires methods and criteria that differ from adult-centered approaches. Lehnert *et al.* (2022) conducted a systematic review of interaction design methods for children and found that evaluation approaches benefit from triangulation — combining self-report instruments with observational and task-based measures — because children's cognitive and expressive abilities vary considerably across age groups. This finding is echoed by Kooijmans *et al.* (2022), who demonstrated that parental or guardian assistance during self-report completion can introduce response bias, underscoring the need for careful procedural controls in child-participant studies. Hagen *et al.* (2023) reviewed digital creativity support tools for children and found that engagement and task completion are consistently shaped by interface simplicity, clarity of guidance, and the responsiveness of feedback — factors that map directly onto Nielsen's learnability and efficiency dimensions. Kakavand *et al.* (2023) evaluated a game-based end-user development tool for children in home settings and reported that usability problems in child-oriented systems tend to cluster around navigation complexity and unclear error recovery, rather than catastrophic technical failures. Lehnert *et al.* (2025) further proposed child-specific usability heuristics for e-assessment tools, arguing that standard adult-oriented heuristics are insufficient for capturing the interaction challenges that children face, particularly in guided learning contexts. Li *et al.* (2025) evaluated an interactive children's education system using human-computer interaction principles and confirmed that visual design clarity, audio support, and response speed are the factors most strongly associated with positive learning outcomes in child-facing applications. Antonelli *et al.* (2023) examined the usability of a transition application for parents of students with visual impairments and found that even well-intentioned educational tools can fail users when interface design does not account for the specific characteristics of the target population — a lesson equally applicable to child-oriented religious learning tools.

A smaller but growing body of work addresses usability in explicitly religious or faith-based application contexts. Ibrahim *et al.* (2023) evaluated the usability of children's learning apps in an Islamic educational context and found that engagement and task completion were strongly influenced by interface simplicity, clarity of guidance, and supportive feedback during errors. Recurring problems included complex navigation, performance delays, and limited onboarding support — particularly in multimedia-heavy applications. Bibi *et al.* (2020) conducted a usability evaluation of Islamic learning mobile applications and reported similar findings, noting that guided onboarding and simplified task flows are particularly important for child users because minor delays or unclear steps can quickly erode attention and cause drop-off. Azahari *et al.* (2025) performed a comparative heuristic assessment of two Islamic mobile applications — Muslim Pro and Sajda — and found that user-centered design principles were inconsistently applied across both platforms, with navigation consistency and error feedback identified as the weakest areas. Allotodang *et al.* (2021) evaluated a Bible learning application using elements of user experience design and reported that visual consistency, clear labeling, and structured content flow were the primary drivers of positive user experience in faith-based learning tools. Taken together, these studies confirm that attractive educational content is insufficient when interaction flows are not learnable or efficient for the target users, and that the gap between content quality and interface quality is particularly consequential in religious learning applications where motivation and comprehension are closely linked. The reviewed literature establishes that usability evaluation of child-oriented applications requires both a validated measurement instrument and methods sensitive to children's developmental characteristics. While prior studies have addressed usability in general educational apps (Hagen *et al.*, 2023; Lehnert *et al.*, 2022), Islamic learning tools (Azahari *et al.*, 2025; Bibi *et al.*, 2020; Ibrahim *et al.*, 2023), and Bible-based applications (Allotodang *et al.*, 2021), no study to date has applied Nielsen's five dimensions with validated psychometric instruments to a child-oriented Christian learning application at scale. The present study addresses this gap by combining a 25-item validated questionnaire with descriptive category analysis across all five usability dimensions, producing a diagnostic profile that can directly inform design decisions for the Superbook Kids Bible App and similar faith-based learning tools.

3. Methodology

This study employed a quantitative approach structured across six sequential stages: problem identification and literature review, questionnaire development, validity and reliability testing, data analysis, formulation of improvement recommendations, and report writing. The complete research workflow is presented in Figure 1.

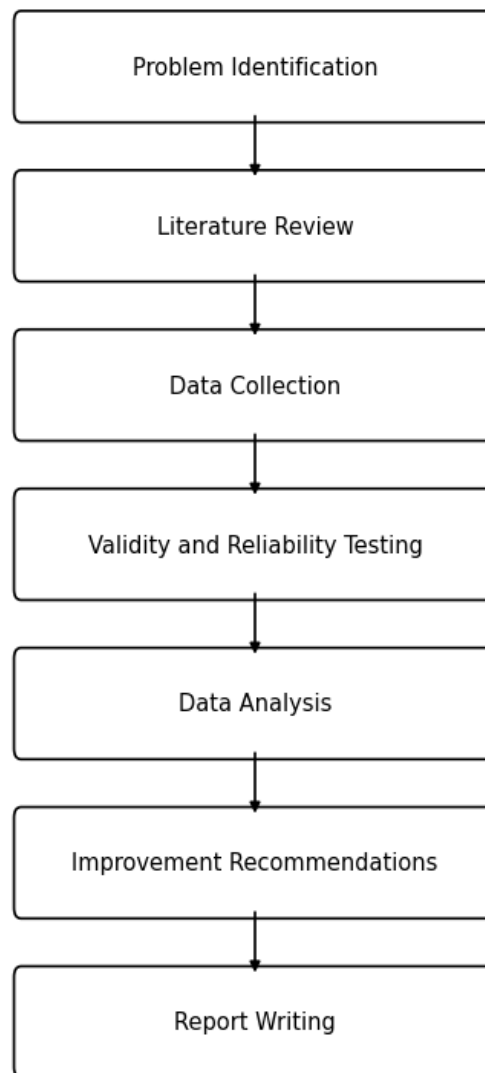


Figure 1. Research Stages.

3.1 Literature Studies

At this stage, a literature review was conducted to collect and synthesize relevant references that support the present study. The research adopts five usability indicators — Learnability, Efficiency, Memorability, Errors, and Satisfaction — as proposed by Jakob Nielsen within the context of Human–Computer Interaction. These indicators are widely used to assess the usability level of information systems (Nielsen, 1993, 2005), and are defined as follows. Learnability evaluates how easily users can learn to operate the system, including their ability to understand the system's purpose and locate the functions or information they need. Efficiency assesses how quickly users can complete tasks once they have learned the system, and identifies potential obstacles or friction points encountered during operation. Memorability measures how well users can re-establish proficiency after a period of not using the system; systems that undergo frequent interface or functional changes may require additional time for users to relearn workflows, potentially increasing repeated errors. Errors examines the frequency and types of errors that occur, including errors caused by the system and those attributable to user actions. Satisfaction reflects users' perceived comfort and positive experience when using the system, indicating the extent to which the system is considered effective, pleasant, and beneficial from the user's perspective.

In addition to using Nielsen's five usability dimensions as the conceptual basis of this study, prior research on child learning applications and faith-based educational apps provides important comparative evidence for interpreting usability issues in religious learning contexts. Existing studies consistently show that children's engagement and task completion are strongly influenced by interface simplicity, clarity of guidance, responsiveness, and supportive feedback during errors (Ibrahim *et al.*, 2023). They also report recurring challenges such as complex navigation, performance delays, feature overload, and limited onboarding support, especially in multimedia-rich applications (Azahari *et al.*, 2025). Evidence from Child-Computer Interaction research further highlights that evaluation methods for children often benefit from triangulation and careful

consideration of children's developmental characteristics, reinforcing the need for child-appropriate usability assessment and design decisions (Lehnert *et al.*, 2022). These findings indicate that attractive educational content alone is insufficient if interaction flows are not efficient and learnable for children. Reviewing related usability evaluations therefore helps position the present work within the broader literature and supports the rationale for applying a systematic dimension-based usability assessment to generate evidence-based design recommendations for child-oriented religious education applications (Allotodang *et al.*, 2021; Bibi *et al.*, 2020). Many studies also emphasize that guided onboarding and simplified task flows are essential for children, because small delays or unclear steps can quickly reduce attention and increase drop-off during learning activities (Bibi *et al.*, 2020; Ibrahim *et al.*, 2023; Lehnert *et al.*, 2025).

3.2 Data Collection

This study applied descriptive quantitative analysis to test and explain the usability variables within a specific usage context. Data were collected through a questionnaire distributed using purposive sampling, where participants were selected based on predefined criteria (Ahmed, 2024). Purposive sampling was used to ensure that respondents had sufficient familiarity and experience with the system being evaluated, thereby improving the relevance and interpretability of the collected data (Andrade, 2021; Campbell *et al.*, 2020; Etikan *et al.*, 2015). The participants were child users aged 6 to 12 years who had used the application at least once per week during the previous four weeks and had interacted with core learning features such as story viewing, quizzes, or games. The inclusion criterion was based on usage frequency rather than purchasing activity, because the study focuses on usability of learning interaction rather than in-app transactions. The questionnaire was developed based on Nielsen's five usability dimensions. Because the participants were children and may have limited reading comprehension, the questionnaire was completed with the assistance of a parent or guardian. Parents or guardians were instructed to read each item aloud, explain unfamiliar words using neutral examples, and record the child's selected Likert-scale response. To minimize parental influence, they were asked not to suggest answers, provide leading interpretations, or respond on behalf of the child. Where possible, children completed the questionnaire immediately after using the application to ensure that responses reflected their direct interaction experience (Kooijmans *et al.*, 2022). As the study involved child participants, standard ethical safeguards were applied. Participation was voluntary, and participants could withdraw at any time without consequences. Written informed consent was obtained from parents or legal guardians prior to data collection, and age-appropriate assent was obtained from the children. Responses were collected anonymously, and no personally identifiable information was recorded. The study activities involved routine app use and were considered minimal risk (Graham *et al.*, 2013). Questionnaire data were processed using descriptive statistical techniques, including standard deviation, minimum score, and maximum score, to capture response dispersion and identify the strongest and weakest usability aspects (Cash *et al.*, 2022). The questionnaire items for each usability variable are listed in Table 1.

Table 1. Questionnaire Items per Usability Dimension

No.	Dimension	Item
1	Learnability	The application is easy to learn and use.
2	Learnability	The instructions in the application are easy to follow.
3	Learnability	The application can be used without assistance from an adult.
4	Learnability	The features in the application are easy to find.
5	Learnability	The font type and text size are easy to read.
6	Learnability	Images and icons are easy to understand.
7	Learnability	The audio can be heard clearly.
8	Learnability	The application provides a help/support feature.
9	Efficiency	Tasks in the application can be completed quickly.
10	Efficiency	Navigation within each menu feels fast.
11	Efficiency	The application's response time is fast.
12	Memorability	The application is easy to remember how to use.
13	Memorability	Menus and functions in the application are easy to remember.
14	Memorability	The application has a consistent design, making it easier to remember.
15	Errors	The application rarely lags or crashes.
16	Errors	Errors rarely occur when using the application.
17	Errors	When an error occurs, it is easy to recover and continue.
18	Errors	The application provides helpful feedback when errors occur.
19	Errors	Any errors that occur do not disrupt gameplay or learning activities.
20	Errors	The application often provides guidance to help avoid repeating the same errors in the future.

21	Satisfaction	The application is enjoyable to use.
22	Satisfaction	The application makes learning more enjoyable.
23	Satisfaction	The application would be recommended to others.
24	Satisfaction	The features in the application are satisfying.
25	Satisfaction	The application's design is comfortable and pleasant to use.

In this study, Slovin's formula was applied to estimate the minimum sample size required to represent the application's user population. The main purpose of using this formula was to obtain data with an acceptable margin of error of 10%, so that the resulting sample could be considered sufficiently representative of the underlying population parameters. Slovin's formula is expressed in Equation (1):

$$n = \frac{N}{1 + Ne^2} \tag{1}$$

Where n is the required sample size, N is the population size, and e is the margin of error. Given an estimated population of N = 10,000,000 users and a margin of error of e = 0.1 (10%), the sample size is calculated as follows:

$$\begin{aligned} n &= \frac{10,000,000}{1 + \frac{(10,000,000 \times 0.1^2)}{10,000,000}} \\ n &= \frac{10,000,000}{1 + (10,000,000 \times 0.01)} \\ n &= \frac{10,000,000}{1 + 100,000} \\ n &= \frac{10,000,000}{100,001} \\ n &= 99.999 \approx 100 \text{ respondents} \end{aligned}$$

Thus, the required sample size is approximately 100 respondents (rounded to the nearest whole number).

3.3 Validity Test

Construct validity was assessed using the Pearson Product–Moment correlation (Karl Pearson), by examining the correlation between each individual item score and the total score of the instrument (item–total correlation) (Kakavand *et al.*, 2023). The Pearson Product–Moment correlation used in this study is shown in Equation (2) (Almanasreh *et al.*, 2019):

$$r = \frac{\sum((X-\bar{X})(Y-\bar{Y}))}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}} \tag{2}$$

Where r is the Pearson correlation coefficient, X and Y are the variables being correlated, and \bar{X} and \bar{Y} are the means of variables X and Y. Based on the validity test results, all questionnaire items were found to be valid, and no items were eliminated. This is indicated by the corrected item–total correlation values, all of which exceeded the critical threshold of 0.271 (Nielsen, 1993). In total, 25 items were confirmed as valid, with validity coefficients ranging from 0.598 to 0.824, indicating moderate to strong item–total relationships.

3.4 Reliability Test

Reliability testing evaluates the consistency and stability of a measurement instrument used in research (Phillips *et al.*, 2021). In other words, it assesses whether the instrument produces consistent results over time and under different conditions, thereby indicating that the measurements are dependable rather than random (Ahmed, 2024). Statistical techniques such as Cronbach's alpha or the intraclass correlation coefficient (ICC) are commonly applied to estimate reliability (Taber, 2018). A higher reliability coefficient indicates that the instrument has stronger internal consistency, which enhances confidence in the credibility and robustness of the study findings. In this study, instrument reliability was examined using Cronbach's alpha, as presented in Equation (3) (Madadzadeh & Bahariniya, 2025):

$$\alpha = \left(\frac{k}{k-1}\right) \left(1 - \left(\frac{\sum \sigma^2 i}{\sigma^2 X}\right)\right) \tag{3}$$

Where α is Cronbach's alpha coefficient, k is the number of items in the scale, $\sigma^2 i$ is the variance of each item, and $\sigma^2 X$ is the variance of the total score. Cronbach's alpha ranges from 0 to 1, where higher values indicate

stronger internal consistency and therefore better reliability of the instrument. Based on the 25 items that were confirmed as valid, the instrument produced a Cronbach's alpha value of 0.754. This result indicates a high level of reliability, suggesting that the questionnaire items measure the usability construct consistently. According to commonly accepted reliability standards, the instrument used to evaluate usability in the application can be classified as reliable (Malkewitz *et al.*, 2023). In practical terms, this implies that the measurement results are stable and dependable, supporting the robustness of the usability assessment.

3.5 Data Analysis

To classify the measurement outcomes for the five usability criteria assessed in this study, each criterion was grouped into three levels: high, moderate, and low. The interval boundaries for these categories were determined using the class-interval approach, calculated by subtracting the minimum value from the maximum value and then dividing the result by the number of categories. The formula used to compute the interval width is shown in Equation (4) (Christanto & Sedyono, 2020):

$$Interval = \frac{Max\ Score - Min\ Score}{Number\ of\ Categories} \quad (4)$$

With 8 valid items for the learnability dimension and a 5-point Likert scale, the maximum possible score is $5 \times 8 = 40$ and the minimum possible score is $1 \times 8 = 8$. Therefore, the interval width is calculated as follows:

$$Interval = \frac{40 - 8}{3}$$

$$Interval = \frac{32}{3} = 10,667 \approx 11$$

Based on these results, the score categories for the Learnability dimension are defined as follows:

- $8 \leq X < 19$: Low (difficult to use)
- $19 \leq X < 30$: Moderate (adequate)
- $30 \leq X \leq 40$: High (easy to use)
- X = total learnability score.

With 3 valid items for the Efficiency dimension and a 5-point Likert scale, the maximum possible score is $5 \times 3 = 15$ and the minimum possible score is $1 \times 3 = 3$. Therefore, the interval width is calculated as follows:

$$Interval = \frac{15 - 3}{3}$$

$$Interval = \frac{12}{3} = 4$$

Based on these results, the score categories for the Efficiency dimension are defined as follows:

- $3 \leq X < 7$: Low (inefficient)
- $7 \leq X < 11$: Moderate (adequate)
- $11 \leq X \leq 15$: High (efficient)
- X = total efficiency score.

With 3 valid items for the Memorability dimension and a 5-point Likert scale, the maximum possible score is $5 \times 3 = 15$ and the minimum possible score is $1 \times 3 = 3$. Therefore, the interval width is calculated as follows:

$$Interval = \frac{15 - 3}{3}$$

$$Interval = \frac{12}{3} = 4$$

Based on these results, the score categories for the Memorability dimension are defined as follows:

- $3 \leq X < 7$: Low (difficult to remember)
- $7 \leq X < 11$: Moderate (adequate)
- $11 \leq X \leq 15$: High (easy to remember)
- X = total memorability score.

With 6 valid items for the Errors dimension and a 5-point Likert scale, the maximum possible score is $5 \times 6 = 30$ and the minimum possible score is $1 \times 6 = 6$. Therefore, the interval width is calculated as follows:

$$\begin{aligned} \text{Interval} &= \frac{30 - 6}{3} \\ \text{Interval} &= \frac{24}{3} = 8 \end{aligned}$$

Based on these results, the score categories for the Errors dimension are defined as follows:

- $6 \leq X < 14$: Low (frequent errors)
- $14 \leq X < 22$: Moderate (adequate)
- $22 \leq X \leq 30$: High (few errors)
- X = total errors score.

With 5 valid items for the Satisfaction dimension and a 5-point Likert scale, the maximum possible score is $5 \times 5 = 25$ and the minimum possible score is $1 \times 5 = 5$. Therefore, the interval width is calculated as follows:

$$\begin{aligned} \text{Interval} &= \frac{25 - 5}{3} \\ \text{Interval} &= \frac{20}{3} = 6,667 \approx 7 \end{aligned}$$

Based on these results, the score categories for the Satisfaction dimension are defined as follows:

- $5 \leq X < 12$: Low (dissatisfied)
- $12 \leq X < 19$: Moderate (adequate)
- $19 \leq X \leq 25$: High (satisfied)
- X = total satisfaction score.

4. Result and Discussion

4.1 Results

The descriptive statistics for the five usability dimensions — Learnability, Efficiency, Memorability, Errors, and Satisfaction — are summarized in Table 2. The Superbook Kids Bible App demonstrates a moderately positive usability profile across the evaluated dimensions, indicating that the application is broadly accessible and engaging for its intended users. The interface layout, visual presentation, and interactive content are generally aligned with child-oriented interaction needs, supporting basic navigation, recognition of icons and multimedia elements, and enjoyment during use. The distribution of scores indicates that many users can complete core activities and interact with the learning features without major difficulty, reflecting an acceptable level of usability for a child-focused educational application. The mean and median values across dimensions further suggest that user perceptions tend to cluster around the moderate range, implying that the application performs reliably for many users but may not yet provide a consistently smooth experience for all children, particularly those who require more guidance or have lower tolerance for delays.

Table 2. Usability Results

	Learnability	Efficiency	Memorability	Errors	Satisfaction
N Valid	100	100	100	100	100
N Missing	0	0	0	0	0
Mean	22.40	8.12	8.65	17.76	14.91
Median	22.00	7.00	8.50	18.00	13.00
Std. Deviation	5.012	2.53	2.757	3.922	4.055
Minimum	15.00	3.00	3.00	6.00	5.00
Maximum	32.00	15.00	15.00	30.00	25.00

The learnability results are presented in Table 3. As the application targets children, learnability is a critical requirement because young users rely heavily on clear cues, simple task flows, and supportive guidance.

Table 3. Learnability Results

Category	Score Range	Frequency (F)	Percentage (%)	Mean	Median	Std. Dev
Low	$8 \leq X < 19$	33	33%	22.40	22.00	5.012
Moderate	$19 \leq X < 30$	51	51%			
High	$30 \leq X \leq 40$	16	16%			

The findings show that 51% of respondents rated learnability as moderate, 16% as high, and 33% as low. This indicates that a substantial proportion of users still encounter difficulties when first learning how to use the application. Efficiency results are shown in Table 4. Efficiency is especially important for children because they typically have shorter attention spans and are more likely to disengage when an application feels slow or requires too many steps.

Table 4. Efficiency Results

Category	Score Range	Frequency (F)	Percentage (%)	Mean	Median	Std. Dev
Low	$3 \leq X < 7$	46	46%	8.12	7.00	2.53
Moderate	$7 \leq X < 11$	42	42%			
High	$11 \leq X \leq 15$	12	12%			

The results indicate that 46% of respondents perceived the application as inefficient, while 42% rated it as moderate, and only 12% rated it as high. This suggests that many users experience friction in completing tasks quickly, possibly due to slow response times, heavy multimedia content, or navigation structures that require multiple actions to reach key features. Memorability results are presented in Table 5. Memorability is important for applications that are used repeatedly, as children often return to preferred learning activities and expect familiar patterns.

Table 5. Memorability Results

Category	Score Range	Frequency (F)	Percentage (%)	Mean	Median	Std. Dev
Low	$3 \leq X < 7$	31	31%	8.65	8.50	2.757
Moderate	$7 \leq X < 11$	37	37%			
High	$11 \leq X \leq 15$	32	32%			

A total of 37% of respondents rated memorability as moderate, 32% as high, and 31% as low. These findings suggest a relatively balanced distribution, indicating that while a portion of users find the application easy to remember, a similar proportion struggles to retain how to navigate and use its features. Error-related results are summarized in Table 6. In usability evaluation, the Errors dimension reflects not only technical stability — such as lag and crashes — but also whether the system helps users recover easily and prevents repeated mistakes.

Table 6. Errors Results

Category	Score Range	Frequency (F)	Percentage (%)	Mean	Median	Std. Dev
Low	$6 \leq X < 14$	14	14%	17.76	18.00	3.922
Moderate	$14 \leq X < 22$	60	60%			
High	$22 \leq X \leq 30$	26	26%			

Most respondents (60%) placed the application in the moderate category for errors, while 26% perceived few errors, and 14% reported frequent errors or disruptions. Although the majority did not experience severe instability, recurring issues such as lag, unclear feedback, or occasional crashes can still interrupt learning and reduce enjoyment. Satisfaction results are presented in Table 7. Satisfaction reflects the overall user experience and is influenced by multiple factors, including usability, performance, visual appeal, and the perceived educational value of the content.

Table 7. Satisfaction Results

Category	Score Range	Frequency (F)	Percentage (%)	Mean	Median	Std. Dev
Low	$5 \leq X < 12$	29	29%	14.91	13.00	4.055
Moderate	$12 \leq X < 19$	37	37%			
High	$19 \leq X \leq 25$	34	34%			

The results indicate that 37% of respondents reported moderate satisfaction, 34% reported high satisfaction, and 29% reported low satisfaction. This distribution suggests that while many users enjoy the application, a

considerable portion remains dissatisfied — likely due to issues associated with efficiency and learnability, as those dimensions show higher proportions in the low category.

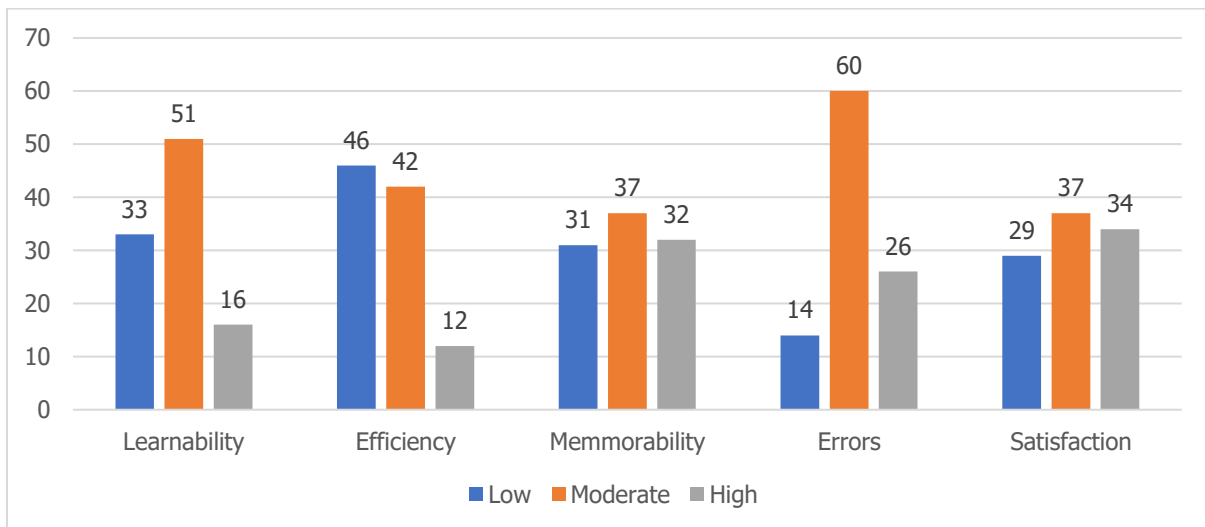


Figure 2. Cross-Dimension Comparison

Figure 2 presents a comparative overview of usability categories across all five dimensions. In the low category, Efficiency shows the highest proportion (46%), indicating that nearly half of respondents perceive limitations in responsiveness and navigation speed. Learnability also has a notable low proportion (33%), suggesting that many children still face challenges during initial use. Memorability (31%) and Satisfaction (29%) further indicate that difficulties in remembering usage patterns and achieving a satisfying experience remain relevant concerns. In contrast, the Errors dimension has the lowest proportion in the low category (14%), implying that severe technical problems are less common than interaction and performance friction. In the moderate category, Errors (60%) and Learnability (51%) dominate, suggesting that the application generally functions without major breakdowns and is reasonably learnable for many users, yet not sufficiently optimized to be perceived as highly efficient or consistently easy to remember. In the high category, Satisfaction (34%) and Memorability (32%) stand out, indicating that a meaningful subset of users feels positive about the overall experience and can recall how to use the application effectively. Learnability (16%) and Efficiency (12%) remain relatively low in the high category, reinforcing the need to improve initial guidance and interaction speed across the user population.

4.2 Discussion

The Superbook Kids Bible App demonstrates acceptable usability, yet it is not fully optimized for child-first interaction. The findings indicate that the application's educational content and engagement features are generally well received, as reflected in the moderate-to-high satisfaction reported by many respondents. However, interaction efficiency and early-stage usability support remain insufficient to ensure consistent adoption across the entire target user group. This is a critical concern for child-oriented applications because children tend to disengage quickly when they experience delays, complex navigation, or unclear system feedback. Improvements in efficiency and learnability are therefore likely to produce the most substantial downstream benefits, particularly in raising overall satisfaction.

Given that efficiency shows the highest proportion in the low category, the first priority should be performance improvement and task-flow simplification. Practical strategies include reducing loading time through asset compression and caching, minimizing heavy animations on initial screens, improving responsiveness by reducing unnecessary background processes, and decreasing the number of taps required to access core learning activities. For child users, rapid access to enjoyable learning content is essential; a shorter and more direct path to key functions such as "start story," "play," or "quiz" is expected to improve perceived speed and reduce early drop-off. The learnability results further indicate the need for stronger guided onboarding, as approximately one-third of respondents reported difficulties in learning how to use the application. To address this, the interface should incorporate a brief, visually driven interactive tutorial during first use, complemented by contextual hints using simple icons and short prompts. Because children's reading abilities vary considerably, learnability can be strengthened by emphasizing visual navigation cues, optional voice instructions, and immediate child-appropriate feedback after actions — for example, confirmation sounds or short animations that indicate a correct step. The consistent placement of primary buttons across screens can also reduce confusion and support faster familiarization.

Although memorability appears relatively balanced, the near-even split between high and low ratings suggests that the application experience is not uniform across users. Improving memorability requires strengthening interface consistency, including standardized iconography, predictable menu structures, and uniform labeling across different sections of the application. Features such as "Recently Used" or "Continue Last Activity" can further support recall while simultaneously reducing navigation steps, enabling children to re-enter familiar content without repeating complex sequences. For the Errors dimension, severe technical failures are not dominant, but the majority of users still placed the application in the moderate category, indicating that error prevention and recovery mechanisms can be improved. Developers should focus on child-appropriate error handling by providing clear, friendly, and actionable messages, along with simple recovery options such as "Try Again" or "Back to Home." Preventive measures — such as larger touch targets, better element spacing, and confirmation prompts when leaving activities — can reduce accidental taps and lower frustration, which in turn supports a more positive user experience.

While satisfaction is already relatively strong for many respondents, it can be further raised by ensuring the overall experience remains smooth, rewarding, and easy to understand from a child's perspective. Satisfaction in child-oriented learning applications is closely linked to how quickly children can access enjoyable content, how clearly the application responds to their actions, and how consistently it maintains a sense of progress. Beyond performance improvements, satisfaction can be strengthened through age-appropriate personalization, stable interactive features, and consistent motivational feedback loops — such as rewards, progress indicators, and completion badges — that reinforce a sense of achievement. Maintaining visual consistency and reducing unpredictable interface changes can also lower discomfort and build trust. Providing clear value indicators such as learning progress summaries, completed story and quiz logs, or simple achievement dashboards can increase perceived usefulness for parents and educators, support guided learning at home or school, and encourage continued engagement over time.

This study has several limitations that should be considered when interpreting the findings. First, the evaluation relied primarily on self-report questionnaire data completed by child participants with parental assistance. Although parents or guardians were instructed to avoid leading explanations and to record children's responses as accurately as possible, this approach may still introduce response bias, including social desirability effects and differences in how items are interpreted across households. Children's cognitive and reading development also varies, which can affect the consistency of self-reported perceptions even when the instrument is valid and reliable. Second, the study did not include direct observational or task-based usability measures, such as task completion time, success rate, navigation path analysis, or objective error frequency. As a result, the findings reflect perceived usability rather than verified performance during real use, and the absence of triangulation limits causal interpretation of why certain dimensions — particularly efficiency — were rated lower. The cross-sectional design captures usability perceptions at a single point in time and cannot account for how usability changes with repeated exposure, learning effects, or application updates. Future work should adopt mixed-method and longitudinal designs by combining questionnaires with structured task observation, usage logging or analytics, and follow-up measurements to provide a more complete and objective understanding of child interaction and sustained engagement.

5. Conclusion

This study concludes that the Superbook Kids Bible App achieves a generally acceptable level of usability for children, but the user experience is not yet consistently optimized across all dimensions. The overall results indicate that the application is perceived as reasonably engaging and supportive for learning, yet the cross-dimension comparison identifies efficiency and learnability as the most prominent constraints. These weaknesses suggest that some children may experience friction during early interaction, including slower response, less streamlined navigation, and limited guidance for understanding how to access and use key features. Although memorability and satisfaction show comparatively stronger outcomes, these strengths can be undermined when performance and onboarding barriers reduce learning continuity and discourage repeated use. In addition, the moderate error profile indicates that stability, feedback clarity, and recovery support should be strengthened to minimize disruptions during learning activities.

This research contributes to the usability and child-computer interaction literature by providing a structured empirical evaluation of a child-oriented religious education application using Nielsen's five usability dimensions, and by operationalizing these dimensions into measurable questionnaire items supported by validity and reliability testing. The study also extends the applicability of usability assessment beyond general educational technology by offering evidence specific to faith-based learning contexts, where engagement, comprehension, and continuity are essential for sustaining learning motivation among children. Furthermore, the comparative analysis across usability categories provides a diagnostic perspective that helps identify which

dimensions act as bottlenecks and how targeted improvements may yield cascading benefits to overall satisfaction and perceived usefulness.

Future research should adopt mixed-method and longitudinal designs to deepen understanding of how children's interaction patterns evolve over time and how usability improvements influence learning outcomes and engagement sustainability. Experimental studies can evaluate the effectiveness of specific interventions such as interactive onboarding, simplified navigation flows, adaptive interfaces by age group, and child-friendly error handling strategies. In addition, future work may explore how gamification mechanisms, personalization, and richer multimodal feedback contribute to efficiency, learnability, and satisfaction, while also considering contextual factors such as device capability, internet connectivity, and parental assistance. Finally, integrating observational task-based usability testing with analytics-driven usage data can provide more objective measures of task completion time, error frequency, and dropout points, enabling stronger causal insights and more generalizable design guidelines for child-oriented religious learning applications.

References

- Ahmed, S. K. (2024). How to choose a sampling technique and determine sample size for research: A simplified guide for researchers. *Oral Oncology Reports, 12*, 100662. <https://doi.org/10.1016/j.oor.2024.100662>
- Allotodang, F., Tolle, H., & Dengen, N. (2021). Design and evaluation of Bible learning application using elements of user experience. *International Journal of Advanced Computer Science and Applications, 12*(5), 422–425. <https://doi.org/10.14569/IJACSA.2021.0120552>
- Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy, 15*(2), 214–221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- Andrade, C. (2021). The inconvenient truth about convenience and purposive samples. *Indian Journal of Psychological Medicine, 43*(1), 86–88. <https://doi.org/10.1177/0253717620977000>
- Antonelli, K., Steverson, A., & Cmar, J. L. (2023). Usability of 4to24: A transition application for parents of students with visual impairments. *Journal of Visual Impairment & Blindness, 117*(1). <https://doi.org/10.1177/0145482X221150239>
- Azahari, N. A., Said, W. M., Hassan, S. N., & Goje, K. (2025). A comparative heuristic assessment of Muslim Pro and Sajda: Evaluating user-centered Islamic mobile applications. *Journal of Hadith Studies, 10*(2), 181–195. <https://doi.org/10.33102/johs.v10i2.401>
- Bibi, S. H., Munaf, R. M., Bawany, N. Z., Shamim, A., & Saleem, Z. (2020). Usability evaluation of Islamic learning mobile applications. *Elkawnie: Journal of Islamic Science and Technology, 8*(1), 1–12. <https://doi.org/10.22373/ekw.v6i1.5920>
- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652–661. <https://doi.org/10.1177/1744987120927206>
- Cash, P., Isaksson, O., Maier, A., & Summers, J. (2022). Sampling in design research: Eight key considerations. *Design Studies, 78*, 101077. <https://doi.org/10.1016/j.destud.2021.101077>
- Casidy, R., Phau, I., & Lwin, M. (2016). The role of religious leaders on digital piracy attitude and intention. *Journal of Retailing and Consumer Services, 32*, 244–252. <https://doi.org/10.1016/j.jretconser.2016.04.006>
- Christanto, H. J., Dewi, C., Sutresno, S. A., & Silalahi, A. D. (2024). Analyzing the use of Chat Generative Pre-Trained Transformer and Artificial Intelligence. *Revue d'Intelligence Artificielle, 38*(4), 1297–1304. <https://doi.org/10.18280/ria.380423>
- Christanto, H. J., & Sedyono, E. (2020). Analisa tingkat usability berdasarkan Human Computer Interaction (HCI) untuk sistem pemesanan tiket online kereta api. *Jurnal Sistem Informasi Bisnis, 10*(2), 163–172. <https://doi.org/10.21456/vol10iss2pp1163-172>

- Christanto, H. J., Sutresno, S. A., Denny, A., & Dewi, C. (2023). Usability analysis of human computer interaction in Google Classroom and Microsoft Teams. *Journal of Theoretical and Applied Information Technology, 101*(16), 6425–6436.
- Christanto, H. J., Sutresno, S. A., Singgalen, Y. A., & Dewi, C. (2024). Analyzing benefits of online train ticket reservation app using technology acceptance. *Ingénierie des Systèmes d'Information, 29*(1), 107–115. <https://doi.org/10.18280/isi.290112>
- El-Asam, A., Katz, A., Street, C., Nazar, N. M., & Livanou, M. (2021). Children's services for the digital age: A qualitative study into current procedures and online risks among service users. *Children and Youth Services Review, 122*, 105872. <https://doi.org/10.1016/j.chidyouth.2020.105872>
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2015). 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>
- Graham, A., Powell, M., Taylor, N., Anderson, D., & Fitzgerald, R. (2013). *Ethical research involving children*. UNICEF Office of Research – Innocenti.
- Hagen, M. H., Cruzes, D. S., Jaccheri, L., & Fails, J. A. (2023). Evaluating digital creativity support for children: A systematic literature review. *International Journal of Child-Computer Interaction, 38*, 100603. <https://doi.org/10.1016/j.ijcci.2023.100603>
- Hiuredhy, D. K., Christanto, H. J., Christine, C., & Sutresno, S. A. (2024). Exploration of modernity: Worship reservation system at Rose of Sharon Church Salatiga utilizing Flutter framework. *Journal of Information Systems and Informatics, 6*(1), 136–152. <https://doi.org/10.51519/journalisi.v6i1.650>
- Ibrahim, A., Al-Rajab, M., Hamid, K., Aqeel, M., Muneer, S., Parveen, M., & Saleem, M. (2023). Usability evaluation of kids' learning apps. In *Proceedings of the 2nd International Conference on Business Analytics for Technology and Security (ICBATS 2023)*. Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ICBATS57792.2023.10111473>
- Itzhak, N. B., Franki, I., Jansen, B., Kostkova, K., Wagemans, J., & Ortibus, E. (2023). Usability and user experience of an individualized and adaptive game-based therapy for children with cerebral visual impairment. *International Journal of Child-Computer Interaction, 35*, 100551. <https://doi.org/10.1016/j.ijcci.2022.100551>
- Kakavand, Z., Shirehjini, A. A., Moghaddam, M. G., & Shirmohammadi, S. (2023). Child-home interaction: Design and usability evaluation of a game-based end-user development for children. *International Journal of Child-Computer Interaction, 37*, 100594. <https://doi.org/10.1016/j.ijcci.2023.100594>
- Kooijmans, R., Langdon, P. E., & Moonen, X. (2022). Assisting children and youth with completing self-report instruments introduces bias: A mixed-method study that includes children and young people's views. *Methods in Psychology, 7*, 100102. <https://doi.org/10.1016/j.metip.2022.100102>
- Lehnert, F. K., Niess, J., Lallemand, C., Markopoulos, P., Fischbach, A., & Koenig, V. (2022). Child-computer interaction: From a systematic review towards an integrated understanding of interaction design methods for children. *International Journal of Child-Computer Interaction, 32*, 100398. <https://doi.org/10.1016/j.ijcci.2021.100398>
- Lehnert, F., Doublet, S., & Sim, G. (2025). Designing usability/user experience heuristics to evaluate e-assessments administered to children. *Computer Standards & Interfaces, 92*, 103933. <https://doi.org/10.1016/j.csi.2024.103933>
- Li, M., Lv, Y., Pu, Y., & Wu, M. (2025). Design and evaluation of children's education interactive learning system based on human computer interaction technology. *Scientific Reports, 15*(1). <https://doi.org/10.1038/s41598-025-90800-y>

- Madadzadeh, F., & Bahariniya, S. (2025). Tutorial on internal consistency assessment by Cronbach's alpha and McDonald's omega. *Perioperative Care and Operating Room Management*, 41, 100568. <https://doi.org/10.1016/j.pcorn.2025.100568>
- Malkewitz, C. P., Schwall, P., Meesters, C., & Hardt, J. (2023). Estimating reliability: A comparison of Cronbach's α , McDonald's ω and the greatest lower bound. *Social Sciences & Humanities Open*, 7(1), 100368. <https://doi.org/10.1016/j.ssaho.2022.100368>
- Mavish, S., Christanto, H. J., & Sutresno, S. A. (2024). Analisis dan redesign aplikasi M-Paspor dengan metode design thinking untuk peningkatan usability dan pengalaman pengguna. *Journal of Information System Research (JOSH)*, 5(4), 947–958. <https://doi.org/10.47065/josh.v5i4.5447>
- Mertala, P., López-Pernas, S., Vartiainen, H., Saqr, M., & Tedre, M. (2024). Digital natives in the scientific literature: A topic modeling approach. *Computers in Human Behavior*, 152, 108076. <https://doi.org/10.1016/j.chb.2023.108076>
- Myers, L. J., & Arterberry, M. E. (2022). Digital media and children under 3 years of age. *Infant Behavior and Development*, 66, 101677.
- Nielsen, J. (1993). *Usability engineering*. Academic Press.
- Nielsen, J. (2005). Usability for the masses. *Journal of Usability Studies*, 1, 2–3. <https://api.semanticscholar.org/CorpusID:63275098>
- Pal, D., & Vanijja, V. (2020). Perceived usability evaluation of Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India. *Children and Youth Services Review*, 119, 105535. <https://doi.org/10.1016/j.childyouth.2020.105535>
- Phillips, S. M., Summerbell, C., Hobbs, M., Hesketh, K. R., Saxena, S., Muir, C., & Hillier-Brown, F. C. (2021). A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children. *International Journal of Behavioral Nutrition and Physical Activity*, 18(1). <https://doi.org/10.1186/s12966-021-01132-9>
- Ridwan, A. (2025, November 25). *Data on Indonesian students using mobile phones, laptops, and the internet in 2025*. Databoks Katadata. <https://databoks.katadata.co.id/en/technology-telecommunications/statistics/69251fd00425f/data-on-indonesian-students-using-mobile-phones-laptops-and-the-internet-in-2025>
- Superbook Kids Bible App [Mobile application]. (2025). The Christian Broadcasting Network (CBN). <https://play.google.com/store/apps/details?id=air.cbn.superbook.bible.app.android&hl=en>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Thompson, P. (2013). The digital natives as learners: Technology use patterns and approaches to learning. *Computers & Education*, 65, 12–33. <https://doi.org/10.1016/j.compedu.2012.12.022>
- Vasalou, A., & Gauthier, A. (2023). The role of CCI in supporting children's engagement with environmental sustainability at a time of climate crisis. *International Journal of Child-Computer Interaction*, 38, 100605. <https://doi.org/10.1016/j.ijcci.2023.100605>