Journal of Instructional and Development Researches
Homepage: https://www.journal.iel-educationorg/index.php/JIDeR
e-ISSN: 2807-5471; p-ISSN: 2807-548X

JIDeR, Vol. 6, No. 5, October 2026
© 2026 Journal of Instructional and Development Researches
Page: 793-804

Development and Content Validation of an Instrument for Evaluating Smart Storybooks with QR Audio

***Salsa Belladinna Putri Utami, Rusilanti, Nur Riska, Leila Siti Chairani, Diffa Putri Arija, Dhea Yuwono Putri**
Universitas Negeri Jakarta, Jakarta, Indonesia
*Email: salsa.belladinna@unj.ac.id (Corresponding Author)

Check for updates OPEN ACCESS

DOI: https://doi.org/10.53621/tdzcac15

Informasi Artikel

Article History:
Received: 18 Juni 2021
Revised: 10 August 2021
Accepted: 15 August 2021
Published: 25 December 2021

Keywords:
Content validity;
Early Childhood Education;
Instrument;
Nutrition Education;
QR Audio;
Storybook.

ABSTRAK

This study aimed to develop an instrument for evaluating smart storybooks with QR audio used in nutrition education for children aged 5–6 years and to report evidence of its content validity. The study used an instrument development and content validation design. Items were drafted from a literature review and a needs analysis, and three experts in educational research and evaluation rated each item on 11 criteria using a five-point scale. Content validity was estimated with Aiken's V, computed for each criterion and averaged per item, with $V \geq 0.800$ as the validity criterion. The instrument comprised 38 items across three domains: content/material (14 items), media (15 items), and language (9 items). Mean Aiken's V was 0.976 for content/material, 0.977 for media, and 0.961 for language, with a weighted mean of 0.973. Item coefficients ranged from 0.856 to 1.000, and all 38 items met the criterion. At the criterion level, 13 of 418 coefficients fell below 0.800, seven of them in two language items, and these low ratings coincided with expert comments on item overlap and on misalignment between indicators and item wording. These results indicate that the instrument is valid and ready for use by educators, researchers, and developers to assess the nutritional content, QR-code access, audio narration, and language of smart storybooks before classroom use. Reliability and agreement among evaluators still need to be tested on different storybook products.

INTRODUCTION

Early childhood is a critical and formative period (Yusuf et al., 2024), as development during this stage can have a significant influence on children's subsequent development into adulthood (Fitriana et al., 2023; Jaoza & S, 2024). The same period shapes positive behaviors (Falensya, 2025), eating habits, and children's understanding of nutrition (Kurniasih & Pratama, 2025), and habits formed then can persist over time (Iwansyah et al., 2026). Nutrition education for young children needs to be delivered through developmentally appropriate learning experiences that are concrete, meaningful, and visually engaging. Learning media help translate nutrition-related concepts into forms that children can understand and relate to their everyday experiences (Varman et al., 2021). Storytelling is one approach used in early childhood learning (Hakim et al., 2026), because it can enhance children's imagination, creativity, and knowledge (Dhea Wandira, 2024). To support storytelling as a learning approach, story-based learning media can be used to provide children with a structured narrative and visual experience. Among such media, storybooks combine narrative and visual elements that can make educational messages more engaging and accessible to young children (Nurfaiza et al., 2025; Putri et al., 2025).

Nutrition is a practical concern in Indonesian early childhood settings. Citing the Indonesian Nutritional Status Survey (SSGI), Falensya (2025) reported that stunting prevalence fell from 21.6% in 2022 to 19.8% in 2024, which remained above the national target of 14%, and concluded that prevention requires cross-sector collaboration and nutrition education from an early age. The same review assigns roles in this task to parents, early childhood education

institutions, health workers, and the community. Classroom observations point in the same direction: in a preliminary study of ten preschool children, Hakim et al. (2026) found that some still brought packaged snacks containing preservatives, sweeteners, and artificial colorings. Teachers who respond to such habits through stories need books whose nutrition messages are correct, and they need a way to check those messages before a book reaches the classroom.

Storybooks are widely used as learning media for young children because they combine narrative and visual elements to present educational messages in an engaging and accessible way (Faizah, 2023). The educational value of storybooks is supported by a systematic review and meta-analysis by Dowdall et al. (2020), which found the potential of storybook reading to support learning during early childhood. In the context of nutrition education, a previous study indicated that storybook-based nutrition education can contribute to improvements in children's nutrition-related knowledge, attitudes, and consumption behavior (Fitryadi et al., 2020). In a systematic review of storytelling for nutrition education, Widiana & Harwanto (2025) reported that storytelling can enhance children's understanding of healthy and balanced diets by engaging their imagination and simplifying complex nutritional information. Research by Nekitsing et al. (2019) demonstrated the potential of storybook-based interventions to influence preschool children's responses to unfamiliar vegetables, providing empirical support for the use of stories in food-related learning. Pradini et al. (2021) also reported the potential of storybooks to support children's nutrition knowledge. Through stories, nutritional concepts can be embedded within characters, situations, and everyday activities, allowing children to encounter health-related messages in a more meaningful context.

The strength of this evidence varies with how and to whom stories are delivered. Dowdall et al. (2020) pooled 19 randomized controlled trials involving 2,594 children aged 1–6 years and found small effects of shared book reading on expressive language (d = 0.41) and receptive language (d = 0.26), and a large effect on caregivers' book-sharing competence (d = 1.01); lower intervention dosage was associated with minimal impact on child language. The nutrition studies cited above mostly involved school-age children. Fitryadi et al. (2020) worked with fifth-grade students, and Pradini et al. (2021) compared 41 fourth-grade students who received an illustrated storybook with 56 controls and found significant gains in knowledge of and attitudes toward vegetables and fruit in the intervention group. Varman et al. (2021) identified nineteen studies conducted in primary schools but six in preschools, and noted that the smaller number of preschool studies and a high risk of bias limit how far the findings generalize. For children aged 5–6 years, the outcome evidence on nutrition storybooks is thin, which places more weight on judging the quality of a book before it is used.

Digital features have expanded storybooks' potential as learning resources (Faizah, 2023). QR codes can provide audio narration that complements printed text and illustrations, potentially enhancing children's engagement and comprehension, especially among pre-readers. Smart storybooks with QR audio therefore combine text, illustrations, and audio in one product. A similar approach appears in Sylvia et al. (2025), who used audiovisual media to introduce healthy foods to young children.

Media of this kind need their own evaluation framework, because a general media checklist does not cover everything a smart storybook has to get right. The content needs to be accurate, relevant, developmentally appropriate (Ermono et al., 2025), and consistent with established nutrition guidance. The media needs to demonstrate appropriate visual design, layout, illustrations, usability, and integration of audio features (Li & Bus, 2023). In addition, the language needs to be clear, simple, effective, and appropriate for the linguistic and cognitive characteristics of children aged 5–6 years. An instrument for these books therefore has to address content, media, and language together.

Previous studies have reported expert validation of storybook media, but the instruments they used were tied to the product under development. Ermono et al. (2025) validated a digital

science storybook through design, language, and material experts and reported validity as percentages, while Arifa and Herwin (2024) reported mean scores from a material expert and a media expert for a digital audio storybook. Hoel and Jernes (2024) proposed seven key strands to help early childhood teachers assess digital books for shared dialogue-based reading, among them verbal text and illustrations, medium flexibility, and interactivity, but the strands do not address nutritional content. Among the studies reviewed here, none offered a scored instrument that combines nutritional accuracy, QR-code access, audio narration, and language suited to children aged 5–6 years. This study therefore aimed to develop an instrument for evaluating smart storybooks with QR audio in early childhood nutrition education and to report evidence of its content validity based on expert judgment. Two questions guided the study: how the domains and items of the instrument were structured, and to what extent the experts rated the items as relevant and appropriate, as indicated by Aiken's V.

METHOD

This study describes the development and content validation of an assessment instrument for smart storybooks with QR audio in early childhood nutrition education. The methodological procedures covered instrument development, expert review, quantitative analysis of content validity, and qualitative examination of expert feedback.

Research Design

This study employed an instrument development and content validation design to establish evidence of content validity for an assessment instrument intended to evaluate smart storybooks with QR audio in early childhood nutrition education. Content validity is the extent to which an instrument adequately represents the construct being measured, as evaluated by qualified experts based on their relevant knowledge and experience (Almanasreh et al., 2019). Aiken's V coefficient provides a quantitative approach to assessing content validity based on expert judgments of assessment items (Kania et al., 2024a). In this study, Aiken's V was used to quantify the content validity of the developed instrument based on ratings provided by three experts in educational research and evaluation. This study focused on the development of assessment items and their evaluation through expert judgment, rather than on testing the effectiveness of the storybook as a learning intervention.

The research process comprised four main stages: (1) specifying the assessment domains and indicators, (2) developing the assessment items, (3) obtaining expert ratings and qualitative feedback, and (4) analyzing the ratings using Aiken's V and reviewing expert comments to inform item refinement. The validation process focused on three domains: content/material, media, and language. Quantitative and qualitative evidence were considered together to evaluate item appropriateness and identify revisions needed to improve the instrument.

Instrument Development

The assessment instrument was developed to evaluate smart storybooks with QR audio for early childhood nutrition education across three domains: content/material, media, and language. The content/material domain addressed the accuracy, relevance, and developmental appropriateness of nutrition-related information. The media domain covered visual design, layout, illustrations, usability, and QR-based audio integration, while the language domain focused on clarity, simplicity, effectiveness, and suitability for children aged 5–6 years.

The assessment items were formulated as declarative statements and rated on a five-point scale, ranging from 1 (lowest rating) to 5 (highest rating). The version submitted for expert review comprised 38 items: 14 in the content/material domain, 15 in the media domain, and 9 in the language domain.

Expert Review and Data Collection

Content validation was conducted through expert judgment involving three university lecturers with expertise in educational research and evaluation. The experts were invited to review the developed assessment instrument to examine the relevance and appropriateness of its items for evaluating smart storybooks with QR audio in early childhood nutrition education. The review focused on whether the items adequately represented the intended assessment domains and whether their statements were sufficiently clear and appropriate for the intended evaluation purposes.

The experts assessed the instrument items using a five-point rating scale: 1 (very poor), 2 (poor), 3 (fair), 4 (good), and 5 (very good). The same three experts reviewed all three domains. Each item was rated against 11 criteria: alignment of the statement with the assessment purpose (P1), with the aspect measured (P2), with the indicator measured (P3), and with the construct measured (P4); relevance of the item content to the aspect assessed (P5); clarity of wording (P6); freedom from ambiguity (P7); consistency of terminology (P8); precision in measuring a single concept (P9); consistency of writing format (P10); and ease of scoring (P11). The criteria were worded for each domain, for example alignment with the purpose of the content, media, or language assessment, so every item received 33 ratings from the three experts. The ratings provided quantitative data for examining item-level content validity, while written comments and suggestions provided qualitative information regarding the strengths and limitations of individual items. These comments were considered to identify statements that might require clarification, more precise wording, or other modifications to improve their relevance and interpretability.

The quantitative ratings and qualitative feedback were subsequently compiled for analysis. The numerical ratings were used to calculate Aiken's V coefficients, while the qualitative feedback was examined to inform decisions regarding item refinement.

Data Analysis

The ratings were analyzed with Aiken's V, an index that ranges from 0 to 1 and is computed as the ratio of the obtained sum of rating differences to the maximum possible sum (Aiken, 1985). For each criterion of an item, $V = \Sigma s / [n(c - 1)]$, where $s = r - lo$, r is the rating assigned by an expert, lo is the lowest category on the scale (1), n is the number of experts (3), and c is the number of rating categories (5). The item coefficient was the mean of its 11 criterion coefficients, which is equivalent to entering all 33 ratings of the item into the same formula. Domain coefficients were the mean of the item coefficients in each domain, and the overall coefficient was weighted by the number of items per domain.

As a worked example, suppose the three experts rate one criterion of an item 5, 5, and 4. Subtracting the lowest category (1) from each rating gives s values of 4, 4, and 3, which add up to $\Sigma s = 11$. The largest possible sum arises when all three experts give the top rating of 5, that is, $3 \times (5 - 1) = 12$. Dividing 11 by 12 gives $V = 0.917$. Following the same steps, ratings of 4, 4, and 4 give $V = 9/12 = 0.750$, and ratings of 4, 4, and 3 give $V = 8/12 = 0.667$. Each point that the experts withhold from the maximum of 12 lowers V by 1/12, or about 0.083. A criterion therefore reaches the 0.800 cut-off only when the three ratings together lose no more than two points, as in ratings of 5, 5, and 4 ($V = 0.917$) or 5, 4, and 4 ($V = 0.833$). The same logic applies to whole items. Each item received 33 ratings (3 experts × 11 criteria), so its maximum sum is $33 \times 4 = 132$. An item coefficient of 0.856, for instance, means that the ratings summed to 113, or 19 points below that maximum.

Coefficients were interpreted with the categories of Retnawati (2016), as cited in, Aljura et al. (2025) which Kania et al. (2024) also adopted: values of 0.400 or below indicate low validity, values above 0.400 up to 0.800 indicate moderate validity, and values above 0.800 indicate high validity. Tajuddin et al. (2025) applied a comparable cut-off, classifying items with $V \geq 0.80$ as

having strong content validity. Items with V of 0.800 or higher were therefore classified as meeting the criterion. Criterion-level coefficients were also examined to locate the aspects on which the experts gave lower ratings. The experts' written comments were sorted by domain and by the kind of change they requested, such as item overlap, unclear wording, or misalignment between indicators and item statements, and these groups informed the revision of item statements. All coefficients refer to the item versions that the experts rated, before revision.

RESULT AND DISCUSSION

Result

The instrument comprised 38 items: 14 in the content/material domain, 15 in the media domain, and 9 in the language domain (Table 1).

Table 1. Distribution of Assessment Items by Domain

No	Assessment Domain	Number of items	Percentage
1	Content/material	14	36.84%
2	Media	15	39.47%
3	Language	9	23.68%
	Total	**38**	**100%**

As shown in Table 1, the media domain contained the largest proportion of items (39.47%), followed by content/material (36.84%) and language (23.68%). This distribution reflects the inclusion of multiple media-related features, particularly the visual presentation and QR-based audio functionality of the storybook.

To illustrate the scope of the instrument, Table 2 presents selected assessment items from the three domains. The examples demonstrate how the assessment dimensions were translated into specific statements concerning nutritional content, media functionality, and language appropriateness.

Table 2. Example Assessment Items by Domain

No	Assessment Domain	Example assessment items
1	Content/material	The benefits of consuming balanced and nutritious food are presented in a way that is easy for children to understand.
2	Content/material	The material presented is consistent with the applicable Balanced Nutrition Guidelines.
3	Media	The QR code can be easily scanned using a compatible device.
4	Media	The narrator's voice is clear, with intonation appropriate for children aged 5–6 years.
5	Language	The sentences use simple structures that are easy for children aged 5–6 years to understand.
6	Language	The language used in the story encourages children to engage in simple thinking through dialogue.

Content validity was examined through the ratings of three experts on 11 criteria for each item. Table 3 summarizes the results. The content/material domain obtained a mean Aiken's V of 0.976, the media domain 0.977, and the language domain 0.961. The weighted mean across the 38 items was 0.973. All 38 items reached the criterion of 0.800. The language domain showed the

widest range of item coefficients (0.856–1.000), whereas the media domain showed the narrowest (0.955–1.000).

Table 3. Summary of Expert Content Validity Ratings

No	Assessment Domain	Number of items	Mean Aiken's V	Minimum–Maximum V	Items meeting the criterion
1	Content/material	14	0.976	0.917–1.000	14 (100%)
2	Media	15	0.977	0.955–1.000	15 (100%)
3	Language	9	0.961	0.856–1.000	9 (100%)
	Total	**38**	**0.973**	**0.856–1.000**	**38 (100%)**

At the criterion level, 405 of the 418 coefficients (11 criteria × 38 items) reached 0.800. The 13 coefficients below this value appeared in six items. In the content or material domain, items 2, 3, and 12 had five criterion coefficients of 0.750 (item 2: P9; item 3: P3 and P6; item 12: P6 and P9). In the media domain, item 3 had one coefficient of 0.750 (P10). The language domain accounted for the remaining seven: item 3 (P3 and P9 = 0.667; P4 = 0.750) and item 5 (P3 and P4 = 0.667; P6 and P9 = 0.750). These two language items also had the lowest item coefficients, 0.886 and 0.856.

Converted back into rating points, the shortfalls were small. Nine of the 13 low coefficients equaled 0.750, a shortfall of three points from the maximum sum of 12 for a criterion, and four equaled 0.667, a shortfall of four points. All four coefficients of 0.667 occurred in language items 3 and 5. The low coefficients clustered on a few criteria: P9 accounted for four of the 13, P3 and P6 for three each, P4 for two, and P10 for one. No criterion coefficient fell below 0.667, so none reached the low-validity band below 0.400. At the item level, language item 5 received 113 of a possible 132 points and language item 3 received 117.

In addition to the numerical ratings, the experts' comments identified aspects requiring clarification or refinement. In the content/material domain, the suggestions concerned overlapping items, the clarity of statements about balanced nutrition, and the need to specify the nutritional guidance used as a reference. In the media domain, the feedback focused on consistency in assessment labels, clearer operational definitions, and more explicit criteria for QR-code readability, audio accessibility, link stability, and device compatibility. In the language domain, the comments emphasized alignment between indicators and item statements, the use of observable behaviors rather than abstract expressions, and the appropriateness of wording for children aged 5–6 years. All 38 items remained in the tabulated instrument; the comments identified opportunities to improve the precision and applicability of individual assessment criteria.

Discussion

Both research questions have direct answers. The instrument organizes 38 items into three domains (content/material, media, and language), and the three experts rated every item at or above the 0.800 criterion, with a weighted mean Aiken's V of 0.973. On this evidence, the instrument is valid in content and can be used to assess smart storybooks with QR audio. This multidimensional approach is consistent with Ermono et al. (2025), who evaluated a digital science storybook in terms of instructional design, language, and content. However, the present study focuses specifically on an evaluation instrument for smart storybooks used in early childhood nutrition education, integrating nutritional content, media functionality, and language appropriateness for children aged 5–6 years.

The media domain accounted for the largest proportion of items (39.47%), reflecting the need to assess visual presentation, usability, and QR-audio functionality. Hoel & Jernes (2024) identified seven quality strands, among them theme, duration, and interactivity, to guide teachers selecting digital picture books for shared reading with young children. Similarly, Arifa & Herwin (2024) evaluated the feasibility, practicality, and effectiveness of a digital audio storybook for

early reading. While these studies examine the quality or educational application of digital storybooks, the present study contributes a structured instrument that explicitly includes QR-code readability, audio accessibility, link stability, and device compatibility in the evaluation of nutrition-focused storybooks.

The audio items rest on experimental evidence about digital picture books. Li and Bus (2023) tested visual and auditory enhancements with 183 children aged 3 to 5 years and found that both kinds of enhancement benefited story comprehension and book-based vocabulary. A version that combined auditory and visual enhancements, however, was less beneficial for comprehension than versions with a single enhancement, particularly in the youngest group. For evaluators, a working QR link and a clear voice are necessary but insufficient: the narration also has to fit the printed text and illustrations without overloading young listeners. The media items on narrator clarity and intonation for children aged 5–6 years address part of this requirement.

The content/material and language domains guard against a different failure: a book whose QR codes scan and whose audio plays can still carry inaccurate or hard-to-follow nutrition messages. Hakim et al. (2026) reported improvements in preschool children's knowledge of healthy foods following a storytelling-based educational activity. Ermono et al. (2025) also incorporated expert assessment of content and language in validating a digital storybook. These studies support the relevance of evaluating both educational content and its presentation. Nevertheless, the present study differs by focusing on the development of assessment criteria rather than examining children's learning outcomes or validating a specific storybook product.

Evidence from nutrition interventions also supports the content/material domain. A systematic review by Varman et al. (2021) found that storybooks, sensory evaluation, games, and creative arts were effective experiential strategies for improving nutrition outcomes in preschool children. In a cluster randomized trial involving 337 children aged 2 to 5 years in 12 UK preschools, Nekitsing et al. (2019) found that children who heard a storybook about the vegetable they were later offered had higher odds of eating it than children who heard a storybook about a different vegetable. Congruence between story content and the foods children encounter appears to matter. Items that check consistency with the Balanced Nutrition Guidelines and the clarity of messages about balanced food give evaluators a way to examine this congruence before a storybook reaches the classroom.

The experts' comments in the present study identified potential item overlap, unclear nutritional statements, insufficiently operationalized media criteria, and issues concerning language-indicator alignment. Reading the coefficients alongside the comments showed where each problem lay. Ahmad et al. (2026), in their validation of a numeracy assessment instrument, also used Aiken's V to analyze expert judgments of item content validity. In the present study, the criterion-level coefficients below 0.800 fell on the items that drew the most specific expert comments. Two experts noted that the indicator for content/material item 2 overlapped with item 1, and two experts noted that the indicators for language items 3 and 5 did not match their statements, for example the indicator wording “critical thinking” against the item wording “simple thinking”. The criteria on which these items lost points fit the same pattern. The four lowest coefficients (0.667) fell on alignment with the indicator (P3), alignment with the construct (P4), and precision in measuring a single concept (P9), the criteria that a mismatch between indicator and statement would affect first. Language item 3, for instance, sits under the indicator “standard terminology” but asks whether the terms suit the language development of children aged 5–6 years. Media item 3 fell below 0.800 only on format consistency (P10), and in the rating form its aspect label carries a letter prefix (“B. Book Cover Design”) that item 2 lacks, which matches the experts' comment on inconsistent assessment labels. The numerical and written evidence therefore pointed to the same revision needs. Further research should examine the instrument's reliability, agreement among evaluators, and applicability across different smart

storybooks.

The coefficients obtained here are higher than those reported in some comparable validations. Kania et al. (2024b) reported Aiken's V values from 0.817 to 0.884 for a mathematical problem-solving instrument rated by five experts, and Ermono et al. (2025) reported percentage validity of 100%, 93%, and 98% from design, language, and content experts for a digital science storybook, with a mean of 97%. Direct comparison is limited because the panels differed in size and the indices summarize different things. Aiken's V averages the distance of each rating from the lowest category, so a panel that rates most items 4 or 5 produces coefficients near the top of the scale, and with three experts the index leaves little room for disagreement to show. Almanasreh et al. (2019) discuss the content validity ratio, the content validity index, and the modified kappa as alternative indices; the modified kappa adjusts for agreement expected by chance. Reporting one of these alongside Aiken's V in a second validation round would let readers judge whether the high coefficients reflect agreement on item quality or general leniency among the raters.

These limitations do not prevent the instrument from being used, but they show where its evidence can be strengthened. All three experts specialized in educational research and evaluation, and the same panel rated all three domains, so their ratings speak most directly to item construction and alignment with the intended domains. Judgments on the accuracy of nutritional content and on the suitability of wording for children aged 5–6 years would gain from a panel that includes a nutrition specialist and an early childhood educator. The panel size also constrains the analysis: with three raters and five categories, a one-point drop in a single rating lowers the coefficient of that criterion by 0.083 and the item coefficient by about 0.008, and the index does not show whether different evaluators would score the same storybook consistently. The coefficients also describe the item versions before revision, so the revised items have not yet been rated. Almanasreh et al. (2019) describe content validation as a sequence of development, judgment and quantification, and revision and reconstruction. The present study covered the first two stages and began the third by documenting expert comments; a second round of expert review on the revised items would complete it.

The next steps can borrow from the sequence that product development studies already use. Arifa and Herwin (2024) moved from expert validation of a digital audio storybook, with mean scores of 3.8 from a material expert and 4.0 from a media expert, to practicality ratings from a teacher (94.44%) and from students (90.83% to 93.89%), and then to a quasi-experimental test of effectiveness. An evaluation instrument needs a different sequence, because its users are evaluators rather than children. After a second expert round, early childhood teachers and storybook developers could apply the instrument to several QR-audio storybooks, and their scores would supply data for estimating inter-rater agreement and internal consistency. Hoel and Jernes (2024) derived their seven strands from 48 video-filmed reading sessions and the teachers' reflections after each session; observing teachers as they apply the instrument could show in the same way which items are hard to use in practice.

In practice, an evaluator can rate each of the 38 statements on the five-point scale, compute domain scores, and return to the lowest-rated items to decide what to revise. A storybook that scores well on media items but poorly on content/material items, for example, would need correction of its nutrition messages before classroom use, regardless of how well its QR audio performs. Developers can also use the domain structure during production, checking the content against the Balanced Nutrition Guidelines before illustration and testing QR codes on several devices before printing.

CONCLUSION

This study developed an instrument for evaluating smart storybooks with QR audio for early childhood nutrition education. The instrument consists of 38 items across three domains:

content/material (14 items), media (15 items), and language (9 items). Based on the expert assessments, the mean Aiken's V coefficients were 0.976 for content/material, 0.977 for media, and 0.961 for language, with an overall weighted mean of 0.973. All 38 items reached the 0.800 criterion, with item coefficients ranging from 0.856 to 1.000. These results show that the instrument is valid in content according to the three experts and is suitable for assessing smart storybooks with QR audio. At the criterion level, 13 of 418 coefficients fell below 0.800. Seven came from language items 3 and 5, including all four coefficients of 0.667, which fell on alignment with the indicator, alignment with the construct, and measurement of a single concept. These criteria correspond to the experts' written comments on mismatches between indicators and item statements. The instrument was designed to address the specific characteristics of smart storybooks used in nutrition education for children aged 5–6 years. In addition to assessing the accuracy and appropriateness of nutritional content, it covers visual presentation, QR-code accessibility, audio quality, and language suitability for young children. Because the domains are scored separately, an evaluator can identify a storybook whose QR audio works well but whose nutrition messages still need correction before classroom use. The instrument can be used by educators, researchers, and developers to review smart storybooks before their use in early childhood education settings. However, the evidence is limited to content validity from a single round of ratings by three experts in educational research and evaluation, and the coefficients describe the item versions before revision. Aiken's V cannot separate agreement on item quality from general leniency among raters, so the high coefficients do not by themselves show that other evaluators would rate the items as favorably.

ACKNOWLEDGMENT

The authors thank the Institute for Research and Community Service (LPPM), Universitas Negeri Jakarta (UNJ), and the Faculty of Engineering, Universitas Negeri Jakarta, for their support of this study. The authors also extend their appreciation to the experts who contributed their time and expertise to the content validation process.

REFERENCE

Ahmad, Santos, M. Dos, Negara, H. R. P., & Sabaryati, J. (2026). Validitas Isi Instrumen Kemampuan Numerasi Siswa Sekolah Dasar Kelas V Menggunakan Indeks Aiken's V. *Jurnal Riset Pendidikan Dasar dan Karakter*, *8*(1), 108–116. https://doi.org/Doi:10.59701/pdk.v8i1.644

Aljura, A. N., Retnawati, H., Dewanti, S. R., Kassymova, G. K., Sotlikova, R., & Septiana, A. R. (2025). Mathematical reasoning and communication word problems with mathematical problem-solving orientation: A relation between the skills. *Journal on Mathematics Education*, *16*(2), 529–558. https://doi.org/10.22342/jme.v16i2.pp529-558

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. In *Research in Social and Administrative Pharmacy* (Vol. 15, Number 2, pp. 214–221). Elsevier Inc. https://doi.org/10.1016/j.sapharm.2018.03.066

Arifa, N., & Herwin, H. (2024). Development of digital audio storybook media to improve early reading skills in first-grade elementary school students. *Jurnal Penelitian Ilmu Pendidikan*, *17*(1), 23–38. https://doi.org/10.21831/jpip.v17i1.91154

Dhea Wandira. (2024). Sosialisasi Story Telling dan Media Poster untuk Menumbuhkan Minat Baca pada Anak Usia Dini di Desa Tanjung Dalam. *ARDHI : Jurnal Pengabdian Dalam Negri*, *2*(4), 23–29. https://doi.org/10.61132/ardhi.v2i4.616

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2020). Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis. In *Child Development* (Vol. 91, Number 2, pp. e383–e399). Blackwell Publishing Inc. https://doi.org/10.1111/cdev.13225

Ermono, E. A. Y., Pritasari, A. C., & Jannah, A. N. (2025). Validasi Buku Cerita Sains Digital Petualangan Respira Berorientasi Perilaku Hidup Bersih dan Sehat. *JUDIKDAS: Jurnal Ilmu Pendidikan Dasar Indonesia*, *4*(2), 120–132. https://doi.org/10.51574/judikdas.v4i2.2304

Faizah, N. (2023). Efektifitas Media Buku Cerita untuk Menambah Kosakata Anak Usia Dini di Kota Jepara. *Journal of Education Research*, *4*(2).

Falensya, R. (2025). Pentingnya Gizi Seimbang dalam Mencegah Stunting pada Anak Usia Dini. *Journal Innovation In Education*, *3*(4), 08–21. https://doi.org/10.59841/inoved.v3i4.3549

Fitriana, D., Jihansyah, I., & Luthfillah, M. (2023). Pentingnya Pendidikan Anak Usia Dini Bagi Tumbuh Kembang Anak. *JCE (Journal of Childhood Education)*, *6*(2), 2598–2184. https://doi.org/10.xxxxx

Fitryadi, A., Asna, A. F., & Noerfitri. (2020). Pengaruh Edukasi Gizi Melalui Media Buku Cerita Terhadap Pengetahuan, Sikap Dan Perilaku Konsumsi Sayur Dan Buah Pada Anak Kelas 5 Sdit Thariq Bin Ziyad Tahun 2019. *Viva Medika: Jurnal Kesehatan, Kebidanan, dan Keperawatan*, *13*(2). http://ejournal.uhb.ac.id/index.php/VM/issue/archive

Hakim, N., Milawati, R., Mamduhah, N., & TestarojaTheresa. (2026). Peningkatan Pengetahuan Anak Usia Dini Mengenai Makanan Sehat Melalui Story Telling di PAUD Ibu Bangsa. *Jurnal Pengabdian Masyarakat*, *5*(1), 225–232. https://doi.org/10.30640/abdimas45.v5i1.5930

Hoel, T., & Jernes, M. (2024). Quality in children's digital picture books: seven key strands for educational reflections for shared dialogue-based reading in early childhood settings. *Early Years*, *44*(3–4), 480–494. https://doi.org/10.1080/09575146.2023.2172552

Iwansyah, A., Zumra, I., Rejki, S., & Muhammadiyah Bima, U. (2026). Pengenalan Makanan Sehat Sejak Dini melalui Metode Bermain dan Media Interaktif. *Pengabdian Dan Pemberdayaan Masyarakat*, *3*(2), 154–162. https://doi.org/10.71301/jp3m.v3i2.266

Jaoza, S. N., & S, A. S. K. (2024). Pentingnya Pendidikan Anak Usia Dini Bagi Tumbuh Kembang Anak. *Global Leadership Organizational Research in Management*, *2*(2), 01–09. https://doi.org/10.59841/glory.v2i2.871

Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024a). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education)*, *10*(1), 64–79. https://doi.org/10.21831/reid.v10i1.71032

Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024b). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education)*, *10*(1), 64–79. https://doi.org/10.21831/reid.v10i1.71032

Kurniasih, S., & Pratama, L. R. (2025). *HAPPY PLATE, PEMBELAJARAN GIZI SEIMBANG ANAK USIA DINI* (Y. Yunita, Ed.). CV. AMY Publishing. https://amypublishing.com/

Li, X., & Bus, A. G. (2023). Efficacy of digital picture book enhancements grounded in multimedia learning principles: Dependent on age? *Learning and Instruction*, *85*. https://doi.org/10.1016/j.learninstruc.2023.101749

Nekitsing, C., Blundell-Birtill, P., Cockroft, J. E., Fildes, A., & Hetherington, M. M. (2019). Increasing Intake of an Unfamiliar Vegetable in Preschool Children Through Learning Using Storybooks and Sensory Play: A Cluster Randomized Trial. *Journal of the Academy of Nutrition and Dietetics*, *119*(12), 2014–2027. https://doi.org/10.1016/j.jand.2019.05.017

Nurfaiza, N., Gusti Bagus Bayu Baruna Ariesta, I., Sagung Intan Pradnyanita, A., & Studi Desain Komunikasi Visual, P. (2025). Perancangan Buku Cerita Ilustrasi Interaktif "Nunu's Little Garden" Untuk Meningkatkan Gemar Makan Sayur Pada Anak Usia 4-6 Tahun Informasi Artikel Abstrak. *Jurnal Selaras Rupa*, *6*(2). https://jurnal.idbbali.ac.id/index.php/selarasrupa

Pradini, I., Angkasa, D., Jusat, I., Purwara Dewanti, L., Wahyuni, Y., Studi Ilmu Gizi, P., Ilmu-Ilmu Kesehatan, F., & Esa Unggul, U. (2021). Pemberian Buku Cerita Bergambar Bertema "Superhero" dapat Meningkatkan Pengetahuan Sayur dan Buah Siswa Sekolah Dasar. In *Jurnal Gizi Unimus* (Vol. 10). http://jurnal.unimus.ac.id

Putri, R. V., Jannah, M., & Meiranny, A. (2025). Literature Review: Edukasi Gizi Seimbang dengan Menggunakan Media Kreatif pada Anak. *J. Midwifery Health Sci. Sultan Agung*, *4*(2), 1–10. https://doi.org/10.30659/jmhsa.v4i2.72

Sylvia, D., Utami, F. B., Rusdiana, N., Alba, M., Daulay, K., Ari, I., Sujera, L., Megawati, S., & Safitri, M. (2025). Jurnal Abdiraja Pengenalan Jenis Makanan Sehat terhadap Anak Usia Dini melalui Media Audio Visual di Kecamatan Pasar Kemis, Kabupaten Tangerang. *Jurnal Abdiraja*, (2).

Tajuddin, N. I. I., Abas, U.-H., Aziz, K. A., Nor, R. N. H., Izni, N. A., Sudin, M. N., Anim, N. A. H. M., & Noor, N. M. (2025). Content Validity Assessment Using Aiken's V: Knowledge Integration Model for Blockchain in Higher Learning Institutions. *IJACSA) International Journal of Advanced Computer Science and Applications*, *16*(6), 601–608. www.ijacsa.thesai.org

Varman, S. D., Cliff, D. P., Jones, R. A., Hammersley, M. L., Zhang, Z., Charlton, K., & Kelly, B. (2021). Experiential learning interventions and healthy eating outcomes in children: A systematic literature review. In *International Journal of Environmental Research and Public Health* (Vol. 18, Number 20). MDPI. https://doi.org/10.3390/ijerph182010824

Widiana, M. E., & Harwanto, F. (2025). The Effectiveness of Storytelling as a Nutrition Education Method in Enhancing the Knowledge of Preschool and Elementary School

Children: A Systematic Review. *Indonesian Journal of Multidiciplinary Research*, *5*(1), 93–102. https://doi.org/10.17509/ijomr.v5i1.81549

Yusuf, O. Y. H., Devi, W. O., Silfana, I., Sunarni, S., Lisnawati, L., Marwah, W. O. Z., Ulfianawati, W. O., Pasa, A., Saharis, H., & Helni, W. O. (2024). Pentingnya Pendidikan Anak Usia Dini. *Ta'rim: Jurnal Pendidikan Dan Anak Usia Dini*, *5*(1), 228–234. https://doi.org/10.59059/tarim.v5i1.1096