Childhood

Volume 1 | Number 1
Published 06-30-2026

Designing an AI-VR Storytelling Animation for Early Reading Stimulation in Early Childhood Education: A Needs Analysis and Design Validation Study

Syamsul Alam Ramli[1*]
[1] Universitas Muhammadiyah Palopo, Indonesia

Abdullah Soares[2]
[2] Universidade Oriental Timor Lorosa'e, Timor-Leste

*Corresponding author: syamsulalamramli@umpalopo.ac.id

Follow this link to access more articles at: https://jurnalpaudindonesia.org/index.php/childhood

Citation
Ramli, S. A. (2026). Designing an AI-VR Storytelling Animation for Early Reading Stimulation in Early Childhood Education: A Needs Analysis and Design Validation Study. Childhood, 1(1), 17-34.
https://jurnalpaudindonesia.org/index.php/childhood/article/view/141

Childhood
Vol. 1, No. 1 (2026) | Pages 17-34

Designing an AI-VR Storytelling Animation for Early Reading Stimulation in Early Childhood Education: A Needs Analysis and Design Validation Study

Syamsul Alam Ramli[1] Universitas Muhammadiyah Palopo, Indonesia
Jacinto de Oliveira Junior[2] Universidade Oriental Timor Lorosa'e, Timor Leste

Abstract

Background: Early reading ability is a fundamental literacy competency for children aged 5-6 years, yet early childhood education (ECE) settings in Indonesia remain predominantly reliant on conventional, non-adaptive instructional media. **Objective:** This study aimed to (1) identify the learning needs and technology readiness of PAUD teachers regarding early reading media, and (2) design a conceptual AI-VR storytelling animation media at Technology Readiness Level (TRL) 1-3. **Methods:** A Research and Development (R&D) approach using the ADDIE model was applied, bounded at TRL 1-3 (Analysis and Design phases). An online, self-administered needs-analysis questionnaire (40 Likert items across five sub-scales, plus open-ended items) and a storyboard response questionnaire (7 Likert items plus open-ended items) were distributed to PAUD teachers via a shared digital form; 167 teachers completed both instruments (97.6% female; 74.3% from Central Sulawesi, 22.2% from South Sulawesi, 3.6% unclassifiable location). A five-episode AI-VR storyboard was rated by three expert validators (two media experts, one ECE curriculum expert) against an 8-item rubric (5-point scale) covering four dimensions; content validity was quantified using Aiken's V (Aiken, 1985). Data were analysed descriptively (mean, SD), for internal consistency (Cronbach's alpha), and for between-region differences (Mann-Whitney U); open-ended responses were examined through keyword-based content analysis. **Results:** Internal consistency was acceptable to excellent across sub-scales (α = 0.74-0.96). Teachers reported strong motivational/innovation needs ($M = 4.39$) and clear prioritisation of early-reading indicators ($M = 4.44$), moderate current-practice limitations ($M = 3.78$) and digital (AI/VR) readiness ($M = 3.82$). No statistically significant differences were found between Central and South Sulawesi respondents on any sub-scale (all $p > 0.05$, Mann-Whitney U). Expert validation of the storyboard yielded a mean Aiken's V of 0.865 (range 0.750-1.000), at the boundary of the acceptable ($V \geq 0.75$) and high ($V \geq 0.87$) validity thresholds. Teacher response to the storyboard was high (overall $M = 4.22$, $SD = 0.63$, $\alpha = 0.95$). **Conclusion:** The AI-VR storytelling animation design is grounded in an empirically documented practitioner needs profile, meets acceptable-to-high expert content validity, and received a positive, internally consistent practitioner response, providing a design-stage foundation for prototype development at TRL 4 and beyond. The self-selected, geographically and educationally skewed nature of the online sample, and the small ($n = 3$) validator panel, are addressed as limitations.

Keywords: Artificial Intelligence, Early Childhood Education, Early Reading, Needs Analysis, Storytelling Animation, Virtual Reality

Introduction

Early reading ability is foundational to academic achievement, with the critical transition from pre-literacy to formal reading-encompassing letter recognition, phoneme awareness, and syllable blending-occurring most intensively between ages 4-7 (OECD, 2023). Indonesia's 2022 PISA data showed only 25% of 15-year-olds met the minimum reading benchmark (OECD, 2023), while adolescent literacy deficits trace directly to inadequate early reading foundations (Laksana et al., 2025), making ECE-phase intervention a strategic priority aligned with SDGs Goal 4 and the RPJMN 2025-2029 digital transformation agenda.

Despite this urgency, Indonesian ECE settings remain dependent on conventional media-flashcards, picture books, teacher-led activities-that are static and non-adaptive (Laksana et al., 2024; Supiani et al., 2025). Teachers report that managing wide reading-readiness variability, delivering individualised instruction, and scaffolding phonological awareness in real time are their most pressing challenges (Demircioğlu, 2025; Karnita et al., 2025; Sugihyono, 2025).

Three research streams offer a convergent solution. Digital storytelling consistently improves phonological awareness and reading motivation in young children (Agusriani et al., 2022; Rini et al., 2024). AI-driven personalisation increases engagement through adaptive difficulty, real-time feedback, and individualised pacing (Alkhawaja et al., 2025; Yanti et al., 2025; Kadir et al., 2025), though design evidence for Indonesian ECE remains limited (Fauziddin & Agustin, 2024). VR's immersive properties enhance engagement and knowledge retention through embodied, contextually grounded experiences (Makransky & Lilleholt, 2018). Bibliometric analysis confirmed that the convergence of AI, VR, and storytelling for early reading in ECE remains globally underexplored and absent from the Indonesian PAUD literature (Shofiyah et al., 2025).

Three critical gaps persist: no integrated AI-VR storytelling design targets early reading indicators for children aged 5-6; empirically documented needs-analysis evidence from Indonesian ECE practitioners at scale is limited; and practitioner-facing design-stage feedback on AI-VR ECE media has rarely been reported quantitatively.

This study addresses these gaps by producing a practitioner-informed AI-VR storytelling animation design at TRL 1-3. Its novelty is threefold: (1) a unified storyboard integrating culturally contextualised narrative, AI adaptive branching logic, and VR immersive environment, targeting empirically identified early reading priorities; (2) a large-sample (N = 167) online needs analysis and design-response dataset from PAUD teachers; and (3) transparent reporting of both the strengths and sampling limitations of practitioner-sourced design evidence. Grounded in Embodied Cognition (Wilson, 2002) and Vygotsky's ZPD (Vygotsky, 1978)-the former informing VR environment design, the latter the AI adaptive logic-the study pursues three objectives: (1) identify PAUD teachers' learning needs and technology readiness; (2) produce an AI-VR storyboard aligned with the PAUD Merdeka Curriculum; and (3) gather practitioner and expert feedback on the storyboard design.

Methods

Research Design

This study employed R&D design (Richey et al., 2003) using the ADDIE model, bounded at TRL 1-3 (Analysis and Design phases). This boundary means no digital prototype was developed and no child learning data were collected; these activities are designated for TRL 4-6 (2027-2029). Table 1 outlines the stage structure. For planning purposes, the Design phase's validation activities were divided into an expert-validation sub-phase (8-item rubric rating by three validators, scored with Aiken's V) and a formative teacher-review sub-phase (online questionnaire distributed to practitioners); no functional digital prototype was built at either sub-phase, so both remain within the TRL 1-3 boundary of storyboard-level design review.

Table 1. ADDIE-Informed Stage Structure and Activities within the TRL 1-3 Scope

Stage	Key Activities	Output	Month
Analysis	Literature review; curriculum analysis; online needs-analysis questionnaire distribution	Needs report; learning-indicator list	1-4
Design - Storyboard Drafting	Storyboard (5 episodes); AI adaptive logic; VR environment concept; instrument development	Storyboard document; AI flowchart; validated instruments	3-6
Design - Expert Validation Sub-phase	Expert rating of storyboard against 8-item rubric by 3 validators (Aiken's V); revision based on feedback	Aiken's V validation report; revised storyboard	6-9
Design - Formative Review Sub-phase	Online storyboard response questionnaire (120+ teachers); quantitative-qualitative data analysis; final design revision; manuscript preparation	Final validated design; research findings; journal manuscript	9-12

Participants and Sampling

Participants were PAUD teachers who voluntarily completed a self-administered online questionnaire (Google Form) shared through regional PAUD teacher networks. This is a convenience/self-selected sample rather than a stratified probabilistic or fixed-quota design; consequently, the achieved sample (N = 167) is geographically and educationally imbalanced. Table 3 presents the full demographic profile. In brief, 74.3% of respondents were classifiable to Central Sulawesi districts (e.g., Donggala, Palu, Toli-Toli, Banggai) and 22.2% to South Sulawesi districts (e.g., Luwu, Palopo, Makassar, Toraja), with 3.6% providing location text that could not be reliably classified. The sample was 97.6% female, predominantly aged 36-45 years (41.3%), and drawn overwhelmingly from private (Swasta) schools (89.8%). Formal education was concentrated at senior-secondary level (SMA, 83.8%), with a minority holding D2 (7.2%), D3 (3.6%), or a non-ECE bachelor's degree (S1 Non-PAUD, 5.4%); no respondent held an ECE-specific S1/S2 qualification. Inclusion required only that respondents currently teach PAUD-aged (5-6-year-old) children; no minimum formal qualification threshold was enforced during recruitment. Three expert validators (two instructional media specialists, one ECE curriculum specialist, all doctoral-qualified) separately reviewed the storyboard design.

Research Instruments

Two respondent-facing instruments were administered, summarised in Table 2. The Needs Analysis Questionnaire comprised 40 five-point Likert items organised into five eight-item sub-scales: (A) Current Practice & Perceived Media Limitations, (B) Motivational & Innovation Needs, (C) Digital (AI/VR) Readiness, (D) Priority Early-Reading Indicators, and (E) AI-VR Concept Acceptance, plus four open-ended questions on teaching difficulties, desired media, expectations of AI-VR storytelling, and other suggestions. The Storyboard Response Questionnaire comprised seven five-point Likert items covering comprehensibility, narrative logic, curricular fit, AI relevance, VR appeal, anticipated engagement, and overall development worthiness, plus three open-ended questions. Internal consistency (Cronbach's alpha) was computed directly from the collected item-level data for both instruments and is reported alongside each sub-scale in the Results (Tables 4 and 7); no separate pilot reliability study was conducted beforehand. The storyboard itself was rated by the three expert validators using an Expert Validation Rubric (8 items; 5-point scale) covering four dimensions: Content & Curriculum Alignment, Narrative & Storytelling Design, AI Adaptive Logic, and VR Environment Design; content validity was quantified per item and overall using Aiken's V, with $V \geq 0.75$ indicating acceptable and $V \geq 0.87$ indicating high validity (Aiken, 1985).

Data Collection Procedure

Data collection proceeded in three phases. (1) Months 1-4: the Needs Analysis Questionnaire was distributed online and completed by 167 respondents; no in-person focus group discussions were conducted, and qualitative insight into teachers' needs was instead drawn from the questionnaire's open-ended items. (2) Months 3-6: the storyboard (five episodes, AI adaptive logic, VR environment concept) was drafted based on the needs-analysis findings. (3) Months 6-9: the three expert validators independently rated the storyboard against the 8-item validation rubric; the storyboard was then distributed, together with the Storyboard Response Questionnaire, to the same pool of respondents (167 completed responses). (4) Months 9-12: data analysis, storyboard refinement, and manuscript preparation.

Data Analysis

Quantitative data were analysed in Python (pandas/SciPy). Descriptive statistics (mean, SD) were computed per item and per sub-scale. Internal consistency was assessed using Cronbach's alpha for each of the five needs-analysis sub-scales and for the storyboard response scale. Because the achieved sample resolved into two substantive regional groups of unequal size (Central Sulawesi n = 124; South Sulawesi n = 37) rather than the originally intended three-region equal-quota design, between-region comparison was conducted using the Mann-Whitney U test (two independent groups) on each sub-scale mean rather than a three-group ANOVA/Kruskal-Wallis test; the six unclassifiable cases were excluded from this specific comparison only, while remaining in all other descriptive analyses. Open-ended responses were examined using keyword-based content tallying by a single analyst to identify recurring response categories; this is a lighter-weight approach than a full multi-phase, multi-coder thematic analysis (cf. Braun & Clarke, 2006), and no inter-coder reliability statistic was computed, which is acknowledged as a limitation.

Ethical Considerations

Ethical approval was obtained from the Institute for Research and Community Service (LPPM) of Universitas Muhammadiyah Palopo. All participants provided written consent after receiving an explanation; the data were anonymized.

Results

Needs Analysis: Teacher Profile and Learning Context

A total of 167 PAUD teachers completed both the Needs Analysis Questionnaire and the Storyboard Response Questionnaire. Table 3 presents the full demographic profile.

Table 3. Demographic Profile of Respondents (N = 167)

Characteristic	Category	n (%), N = 167
Gender	Female / Male	163 (97.6%) / 4 (2.4%)
Age	<25 / 25-35 / 36-45 / >45 years	14 (8.4%) / 63 (37.7%) / 69 (41.3%) / 21 (12.6%)
Formal education	Senior secondary (SMA) / D2 / D3 / S1 (non-ECE)	140 (83.8%) / 12 (7.2%) / 6 (3.6%) / 9 (5.4%)

Characteristic	Category	n (%), N = 167
Teaching experience	<3 / 3-5 / 6-10 / >10 years	7 (4.2%) / 35 (21.0%) / 62 (37.1%) / 63 (37.7%)
School status	Private (Swasta) / Public (Negeri)	150 (89.8%) / 17 (10.2%)
Self-reported region (classified by province)	Central Sulawesi / South Sulawesi / Unclassifiable	124 (74.3%) / 37 (22.2%) / 6 (3.6%)

The sample was almost entirely female (97.6%) and drawn predominantly from private ECE institutions (89.8%). Most respondents (78.8%) held 6 or more years of teaching experience, indicating a practically experienced respondent base, despite the fact that formal ECE-specific qualifications were rare: 83.8% reported senior-secondary (SMA) as their highest formal education, and no respondent reported an ECE-specific tertiary degree. This qualification profile should be kept in mind when interpreting the needs-analysis findings, as it reflects the realities of front-line PAUD staffing in the sampled regions rather than a credentialed specialist sub-population.

Needs Analysis: Questionnaire Results

Table 4 presents item- and sub-scale-level results for the 40-item Needs Analysis Questionnaire, with Cronbach's alpha reported per sub-scale.

Table 4. Needs Analysis Questionnaire Results by Sub-scale (N = 167)

Item	Mean	SD	Category
A. Current Practice & Perceived Media Limitations (Cronbach's α = 0.74)			
I still use printed material as the main medium for reading instruction.	3.89	0.89	High/Agree
The media currently available already meets the needs of all children.	3.56	1.05	High/Agree
Children easily lose attention during reading instruction.	3.53	1.05	High/Agree
I find it difficult to teach children with varying reading abilities.	3.42	1.08	High/Agree
The media I use has not been able to meet each child's needs.	3.47	1.06	High/Agree
Instruction is still dominated by conventional methods.	3.56	0.88	High/Agree
I need a wider variety of more engaging media.	4.43	0.76	Very High/Strongly Agree
Reading instruction requires more interactive media.	4.35	0.79	Very High/Strongly Agree
Sub-scale mean	**3.78**	**0.57**	-
B. Motivational & Innovation Needs (Cronbach's α = 0.91)			
I need media that increases children's learning motivation.	4.51	0.76	Very High/Strongly Agree

Item	Mean	SD	Category
Media should adapt to each child's ability.	4.28	0.86	Very High/Strongly Agree
Engaging media increases children's concentration.	4.53	0.70	Very High/Strongly Agree
Storytelling is more effective than conventional instruction.	4.07	0.79	High/Agree
Animation makes children more interested in learning to read.	4.38	0.77	Very High/Strongly Agree
Reading instruction should be delivered through play.	4.37	0.86	Very High/Strongly Agree
Media should provide an enjoyable learning experience.	4.55	0.65	Very High/Strongly Agree
I need innovative media that keeps pace with technological development.	4.42	0.65	Very High/Strongly Agree
Sub-scale mean	**4.39**	**0.60**	**-**
C. Digital (AI/VR) Readiness (Cronbach's α = 0.82)			
I am familiar with the concept of Artificial Intelligence (AI).	3.73	0.85	High/Agree
I am familiar with the concept of Virtual Reality (VR).	3.31	0.88	Moderate
I have participated in digital-media training.	3.23	1.12	Moderate
I am able to use digital devices.	3.72	0.87	High/Agree
I am willing to learn about AI.	4.32	0.73	Very High/Strongly Agree
I am willing to learn about VR.	4.25	0.77	Very High/Strongly Agree
My school has digital media facilities.	3.72	1.17	High/Agree
I am ready to use new technology.	4.25	0.78	Very High/Strongly Agree
Sub-scale mean	**3.82**	**0.60**	**-**
D. Priority Early-Reading Indicators (Cronbach's α = 0.95)			
Letter recognition is very important.	4.57	0.74	Very High/Strongly Agree
Recognising letter sounds (phonemes) is important.	4.53	0.76	Very High/Strongly Agree
Syllable-blending practice needs to be conducted gradually.	4.44	0.74	Very High/Strongly Agree
Children need to be able to read simple words.	4.29	0.85	Very High/Strongly Agree
Reading instruction should be delivered through play.	4.36	0.84	Very High/Strongly Agree
Picture books help children learn to read.	4.49	0.73	Very High/Strongly Agree
Visual media helps children recognise letters.	4.41	0.74	Very High/Strongly Agree
Instruction must align with children's developmental stage.	4.46	0.78	Very High/Strongly Agree
Sub-scale mean	**4.44**	**0.67**	**-**
E. AI-VR Concept Acceptance (Cronbach's α = 0.96)			

Item	Mean	SD	Category
I am interested in using AI-VR Storytelling once it is developed.	4.14	0.84	High/Agree
I believe this concept can increase learning motivation.	4.23	0.78	Very High/Strongly Agree
I believe this concept helps early reading instruction.	4.16	0.74	High/Agree
I am willing to participate in training on this media.	4.32	0.74	Very High/Strongly Agree
I will use this media once it is available.	4.28	0.76	Very High/Strongly Agree
I would recommend this media to colleagues.	4.16	0.70	High/Agree
I support the development of AI-VR learning media.	4.25	0.70	Very High/Strongly Agree
This concept has potential for implementation in PAUD settings.	4.25	0.73	Very High/Strongly Agree
Sub-scale mean	**4.22**	**0.67**	**-**

Internal consistency ranged from acceptable (Sub-scale A, $\alpha = 0.74$) to excellent (Sub-scale E, $\alpha = 0.96$). Sub-scale B (Motivational & Innovation Needs, M = 4.39) and Sub-scale D (Priority Early-Reading Indicators, M = 4.44) received the highest agreement, indicating strong practitioner-felt demand for more engaging, adaptive media and clear consensus on which early-reading skills to prioritise (letter recognition M = 4.57; phoneme recognition M = 4.53; syllable blending M = 4.44). Sub-scale A (Current Practice & Perceived Media Limitations, M = 3.78) and Sub-scale C (Digital (AI/VR) Readiness, M = 3.82) were more moderate, reflecting a teacher population that is experienced in daily practice but only moderately exposed to digital or immersive technology (e.g., prior digital-media training M = 3.23; VR familiarity M = 3.31). One item in Sub-scale A ('The media currently available already meets the needs of all children', M = 3.56) was worded in the opposite direction from the rest of the sub-scale; reverse-scoring this item during a sensitivity check reduced rather than improved sub-scale alpha (0.64 vs 0.74 as administered), suggesting the item may not have functioned as a true reverse-worded check in this sample-a wording ambiguity noted in the Limitations.

Between-region comparison (Table 8) found no statistically significant difference between Central Sulawesi (n = 124) and South Sulawesi (n = 37) respondents on any of the five needs-analysis sub-scales or the storyboard response scale (all $p > 0.05$, Mann-Whitney U), suggesting a broadly similar needs and response profile across the two regional groups represented in this sample, despite their unequal size.

Table 8. Comparison of Sub-scale Scores between Regional Groups (Mann-Whitney U)

Construct	Central Sulawesi (n=124) M (SD)	South Sulawesi (n=37) M (SD)	Mann-Whitney U	p	Sig.
A. Current Practice & Media Limitations	3.81 (0.49)	3.69 (0.74)	2374.0	0.748	ns
B. Motivational & Innovation Needs	4.41 (0.45)	4.32 (0.90)	2148.5	0.557	ns
C. Digital (AI/VR) Readiness	3.80 (0.52)	3.89 (0.82)	1881.5	0.097	ns
D. Priority Early-Reading Indicators	4.49 (0.47)	4.31 (1.03)	2238.0	0.821	ns
E. AI-VR Concept Acceptance	4.24 (0.53)	4.17 (0.98)	2085.0	0.388	ns
F. Storyboard Response	4.24 (0.60)	4.16 (0.74)	2289.5	0.987	ns

Qualitative Content Analysis of Open-Ended Responses

Keyword-based tallying of the open-ended question on the greatest difficulty in teaching early reading (n = 167 responses) identified children not yet recognising letters as the most frequently mentioned difficulty (57.5%), followed by low attention/concentration (44.3%), low reading motivation (31.7%), diverse ability levels within a class (16.2%), and limited media/facilities (13.2%). For the question on the media most needed, visual/colourful media was most frequently mentioned (59.3%), followed by interactive media (34.7%), digital/technology-based media (28.7%), play-based media (25.7%), and storytelling-based media (25.1%). These patterns are consistent with, and help explain, the quantitative sub-scale findings above (high scores for engaging/adaptive media need; moderate digital-technology familiarity).

Storyboard Design

Based on the needs-analysis findings-particularly the priority placed on letter recognition, syllable blending, and simple-word reading, and the expressed preference for visual, interactive, story-based media-the Design phase produced a five-episode AI-VR storytelling animation storyboard. Table 5 presents the design specification for each episode.

Table 5. AI-VR Storytelling Animation: Storyboard Episode Specifications

Episode	Story Title	Setting	AI Adaptive Logic	Early Reading Target
1	The Letter Keeper's Forest	Enchanted forest	Phoneme response speed → adjusts hint frequency	Letter-sound correspondence (A-E)
2	Underwater Alphabet Adventure	Coral reef world	Syllable accuracy → adjusts difficulty level	Onset-rime syllable blending
3	The Village Word Market	Traditional village market	Word-inference accuracy → adjusts narration pace	Simple word reading in context
4	Space Station Spell-Out	Outer space station	Multi-letter pattern accuracy → adjusts scaffold level	CVC word formation
5	Harvest Day Story	Rural harvest scene	Integrated score → adjusts episode complexity	Short sentence comprehension

Each episode follows Gagné's nine events of instruction (Mayer, 2009) adapted for early childhood: attention gained through the VR opening scene; objective communicated via narrative character invitation; prior knowledge stimulated through a brief review; stimulus presented through the main storytelling sequence; guidance provided through AI-delivered hints; performance elicited through interactive activities; immediate audio-visual feedback provided; performance assessed through the AI accuracy module; and retention enhanced through end-of-episode summary narration.

Expert Validation Results

The storyboard design was rated by three expert validators (two media specialists, one ECE curriculum specialist) against the 8-item rubric. Content validity was quantified using Aiken's V (Aiken, 1985). Table 6 presents the results.

Table 6. Expert Validation Results - Storyboard Design (Aiken's V)

Dimension	Item	Expert 1 (Media)	Expert 2 (Media)	Expert 3 (ECE)	Aiken's V	Category
Content & Curriculum	Alignment of storyboard content with the PAUD Merdeka Curriculum	5	5	5	1.000	High
Content & Curriculum	Clarity and sequencing of early-reading learning objectives	4	5	5	0.917	High
Narrative & Storytelling	Appropriateness of story themes for children aged 5-6 years	4	4	5	0.833	Acceptable
Narrative & Storytelling	Clarity and flow of the narrative script (dialogue, narration)	4	5	4	0.833	Acceptable
AI Adaptive Logic	Logical coherence of the AI adaptive branching scenarios	4	5	5	0.917	High

Dimension	Item	Expert 1 (Media)	Expert 2 (Media)	Expert 3 (ECE)	Aiken's V	Category
AI Adaptive Logic	Suitability of the AI difficulty-adjustment logic for young learners	4	4	5	0.833	Acceptable
VR Environment Design	Child-appropriateness and safety of the planned VR visual environments	4	4	4	0.750	Acceptable
VR Environment Design	Technical feasibility and usability of the proposed VR interaction design	4	4	5	0.833	Acceptable
Overall	**Mean (n = 3 validators, 8 items)**	**4.00**	**4.50**	**4.63**	**0.865**	**Acceptable**

All eight items yielded Aiken's V values between 0.750 and 1.000 (mean = 0.865), at the boundary between the acceptable ($V \geq 0.75$) and high ($V \geq 0.87$) validity thresholds (Aiken, 1985). The highest index was obtained for alignment of storyboard content with the PAUD Merdeka Curriculum (V = 1.000, unanimous top rating from all three validators). The two Content & Curriculum items achieved the highest sub-mean (0.958), followed by AI Adaptive Logic (0.875), Narrative & Storytelling Design (0.833), and VR Environment Design (0.792, the lowest of the four dimensions). No item fell below the acceptable threshold. Given the small validator panel (n = 3), these coefficients should be interpreted as an initial, though genuine, content-validity estimate rather than a definitive one; a larger or more diverse validator panel is recommended before TRL 4 prototyping.

Teacher Response to Storyboard Design

Following expert review, all 167 respondents reviewed the storyboard and completed the Storyboard Response Questionnaire. Table 7 presents item- and scale-level results.

Table 7. Teacher Response to the AI-VR Storytelling Animation Storyboard (N = 167)

Response Item	Mean	SD	Category
The storyboard is easy to understand.	4.16	0.74	High/Agree
The storyline is logically structured.	4.20	0.68	Very High
The early-reading content is appropriate for children aged 5-6 years.	4.23	0.70	Very High
The AI concept in the storyboard is relevant for reading instruction.	4.18	0.75	High/Agree
The planned VR environment is appealing and appropriate for young children.	4.19	0.76	High/Agree
The storyboard has potential to increase children's engagement in reading instruction.	4.31	0.68	Very High
Overall, this storyboard is worth developing into an AI-VR based learning medium.	4.30	0.69	Very High
Overall Mean (Cronbach's α = 0.95)	**4.22**	**0.63**	**Very High**

The scale showed excellent internal consistency (α = 0.95) and an overall mean of 4.22 (SD = 0.63), indicating generally high, consistent practitioner endorsement. The highest-rated items were the storyboard's potential to increase children's engagement (M = 4.31) and its overall worthiness for development (M = 4.30); the comprehensibility item received the lowest, though still favourable, mean (M = 4.16). Keyword tallying of the open-ended question on aspects needing improvement (n = 167) found that 29.9% of respondents indicated no changes were needed, while the remainder most frequently raised visual/animation consistency (31.7%), narrative/text simplification (19.2%), and scene pacing/tempo (16.2%) as areas for refinement-feedback that has been incorporated into the storyboard revision plan for the next design iteration.

Discussion

Needs Analysis: Empirical Design Grounding

The relatively moderate scores on current-practice limitations (M = 3.78) and digital readiness (M = 3.82), alongside strong motivational/innovation-need scores (M = 4.39) and clear reading-priority consensus (M = 4.44), depict a teacher population that is experienced and pedagogically engaged but only moderately exposed to digital or immersive technology-consistent with the broader pattern of conventional-media reliance documented across Indonesian ECE settings (Supiani et al., 2025; Fauziddin & Agustin, 2024). This profile differs somewhat from an idealised technology-ready adopter population, and the moderate (rather than uniformly high) technology-readiness scores are arguably a more realistic empirical picture than a design brief based on assumption alone. The consistently high priority placed on letter recognition, phoneme awareness, and syllable blending directly informed each episode's activity structure, distinguishing this study from prior AI-VR research driven primarily by technological capability rather than documented practitioner need (Tsang et al., 2025; Huang et al., 2024). The absence of significant differences between Central and South Sulawesi respondents (Table 8) suggests that, within this sample, these needs and readiness patterns are not strongly region-specific-though the unequal group sizes and non-probabilistic sampling caution against generalising this finding to the wider PAUD teacher population (Karnita et al., 2025). The open-ended content-analysis themes-letter unfamiliarity, attention/concentration difficulty, and a preference for visual, story-based media-reinforce and help explain the quantitative sub-scale patterns, and support the locally contextualised, story-driven design approach adopted here (Rini et al., 2024; Tarigan et al., 2024).

Design Validity: What the Expert Review and Teacher Response Tell Us

The mean Aiken's V of 0.865 sits just below the high-validity threshold (V ≥ 0.87) but comfortably above the acceptable threshold (V ≥ 0.75) (Aiken, 1985), indicating a design that domain experts judged sound overall while identifying room for refinement, particularly in the VR Environment Design dimension (sub-mean = 0.792)-consistent with documented challenges in designing head-mounted or immersive VR experiences for young children (Makransky & Lilleholt, 2018). The unanimous top rating for curriculum alignment (V = 1.000) suggests the storyboard's instructional content is well anchored to the PAUD Merdeka Curriculum, which is arguably the most safety- and standards-critical of the four validated dimensions. Because the validator panel comprised only three experts, these coefficients are best read as an initial content-validity estimate rather than a definitive one. This expert evidence converges with, and is reinforced by, the considerably larger teacher-response evidence base: an overall mean of 4.22 (SD = 0.63, α = 0.95) across 167 practitioners represents independent practitioner validity (Nieveen & Folmer, 2013) from a much larger and more systematically measured source than the three-validator expert panel alone. The high internal consistency of both the needs-analysis and storyboard-response scales indicates that, whatever their sampling limitations, respondents answered these items in a coherent, non-random pattern-supporting the basic measurement quality of the instruments even though a priori pilot reliability testing was not conducted.

Teacher Response: Practitioner Validity

The overall practitioner response mean of 4.22 (Very High-range) from 167 PAUD teachers constitutes a substantive, empirically grounded practitioner-validity signal. The highest-rated items-engagement potential (M = 4.31) and overall development worthiness (M = 4.30)-are encouraging, as practitioner endorsement precedes successful ECE technology adoption in the literature (Demircioğlu, 2025; Karnita et al., 2025). The comparatively lower (though still favourable) comprehensibility score (M = 4.16), together with qualitative feedback on visual consistency, narrative simplification, and scene pacing, indicates specific, actionable refinements for the next storyboard iteration and for explanatory/training materials to accompany TRL 4 prototype development (Kadir et al., 2025; Fauziddin & Agustin, 2024).

Limitations and Directions for TKT 4-6

Several limitations apply. First, the TRL 1-3 boundary precludes prototype testing and child learning data; effectiveness evidence is reserved for TRL 4-6 (2027-2029). Second, the sample was self-selected via an openly shared online form rather than drawn through stratified or fixed-quota sampling; it is geographically skewed toward Central Sulawesi

(74.3%) rather than evenly distributed across the originally intended three South Sulawesi regions, and it is heavily skewed toward private-school (89.8%) and non-ECE-credentialed (83.8% senior-secondary-only) respondents. Findings should therefore be interpreted as reflecting this specific practitioner sample rather than as a representative picture of Indonesian PAUD teachers generally, and formal generalisation claims are not warranted. Third, although expert content validity was quantified using Aiken's V (mean = 0.865, range 0.750-1.000), the validator panel comprised only three experts, which is standard practice for Aiken's V in Indonesian R&D research (Aiken, 1985) but limits inter-rater diversity; a larger and/or more diverse panel would strengthen confidence in these coefficients. Fourth, although a genuine between-region statistical comparison was conducted (Mann-Whitney U, Table 8), the two compared groups were of markedly unequal size (n = 124 vs 37) and were derived post hoc from free-text location fields rather than a planned equal-quota design, which limits the precision and generalisability of this comparison. Fifth, no in-person focus group discussions were conducted; qualitative insight was instead drawn from open-ended questionnaire items, analysed through single-analyst keyword tallying rather than a full multi-phase, multi-coder thematic analysis, and no inter-coder reliability statistic was computed. Sixth, one needs-analysis item (Sub-scale A, item 2) showed evidence of not functioning as intended when reverse-scored, indicating a possible wording-clarity issue that should be revised before the instrument is reused. Finally, storyboard responses reflect appraisal of a static/printed design document rather than an interactive prototype.

Conclusion

This study conducted a large-sample (N = 167) online needs analysis and produced an AI-VR storytelling animation storyboard for early reading stimulation in children aged 5-6 years, at TRL 1-3 within an R&D ADDIE framework. Three conclusions are drawn.

Needs-analysis data from 167 self-selected PAUD teachers, predominantly from Central Sulawesi and South Sulawesi, showed internally consistent (α = 0.74-0.96), moderate-to-high demand for more engaging, adaptive reading media (Sub-scale B, M = 4.39) and clear practitioner consensus on priority early-reading indicators-letter recognition, phoneme awareness, and syllable blending (Sub-scale D, M = 4.44)-alongside only moderate current digital/AI-VR readiness (Sub-scale C, M = 3.82). No statistically significant differences were found between the two regional groups represented in the sample (all $p > 0.05$, Mann-Whitney U).

The five-episode AI-VR storyboard-integrating culturally contextualised settings, Gagné-structured sequences, AI adaptive branching logic, and VR immersive environments-was directly informed by these needs-analysis findings and achieved acceptable-to-high expert content validity (mean Aiken's V = 0.865; range 0.750-1.000) from three expert validators.

Teacher response from all 167 practitioners yielded a high, internally consistent overall mean of 4.22 ($SD = 0.63$, $\alpha = 0.95$), with the strongest agreement on the storyboard's potential to increase children's engagement and its overall worthiness for further development. Together, these findings provide a practitioner-grounded, though sample-limited, design-stage evidence base for prototype development and structured expert and effectiveness testing at TRL 4 (2027-2029), pending the sampling, validation, and analytic improvements identified above as limitations.

Acknowledgements

The authors gratefully acknowledge the support of the Early Childhood Education (PGPAUD) Study Program at Muhammadiyah University of Palopo. They express their sincere appreciation to the early childhood education (PAUD) schools in Central Sulawesi and South Sulawesi for their time, insight, and willingness to participate in this research.

Declaration on the Use of Artificial Intelligence

The authors declare that artificial intelligence (AI) tools were used in the preparation of this manuscript. Claude.ai (Anthropic) was used solely for language editing and grammatical improvement. All ideas, interpretations, data analyses, and conclusions presented in this manuscript were developed independently by the authors. The authors assume full responsibility for the accuracy, integrity, and originality of the entire content of this work.

References

Adiningsih, Y., Royani, I., & Latipah, D. (2024). Edukasi dan pendampingan pembuatan bahan ajar untuk anak usia dini berbasis artificial intelligence (AI). *Jurnal SOLMA*, 13(3), 2762-2774. https://doi.org/10.22236/solma.v13i3.16379

Agusriani, A., Sumiati, S., Ismail, W., Nurhayati, A., & Rachmatiah, S. (2022). Penggunaan alat peraga dalam metode bercerita untuk meningkatkan perbendaharaan kata anak 5-6 tahun. *KHIDMAH: Jurnal Pengabdian kepada Masyarakat,* 2(2), 141-150. https://doi.org/10.24252/khidmah.v2i2.30214

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement,* 45(1), 131-142. https://doi.org/10.1177/0013164485451012

Alkhawaja, L., Idris, M., Al-Sayyed, S. W., & Jaber, A. R. (2025). Exploring the impact of artificial intelligence on students' skills for sustainable development in education. *Frontiers in Education,* 10. https://doi.org/10.3389/feduc.2025.1691148

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

Demircioğlu, D. (2025). Child education and artificial intelligence in the future. *DergiPark (Istanbul University)*. https://dergipark.org.tr/en/pub/kitod/issue/90356/1619610

Fauziddin, M., & Agustin, M. (2024). Semantic literature review: Manfaat artificial intelligence (AI) pada pendidikan anak usia dini di Indonesia. *Jurnal Obsesi*, 8(6), 1475-1488. https://doi.org/10.31004/obsesi.v8i6.6236

Huang, R., Bao, Y., Wang, Y., & Wei, J. (2024). Exploring AI tools in early childhood education: Usage patterns, functions, and developmental outcomes. *In Education and Human Development*. https://doi.org/10.5772/intechopen.1007116

Kadir, A., Sairin, S., Madjid, A., Widigdo, M. S. A., Ramli, S. A., & Asrifan, A. (2025). Creating holistic learning environments: Integrating AI, education, and psychological principles. In Human-Centered Learning Design in the AI Era (pp. 29-59*). IGI Global*. https://doi.org/10.4018/979-8-3373-5786-7.ch002

Karnita, A., Aisyah, S., Mustapa, N., Syarah, E. S., Jovanka, D. R., & Mahamud, S. bt. (2025). Exploring the impact of digital literacy on cognitive development in early childhood education: A systematic literature review. *Al-Athfal: Jurnal Pendidikan Anak*, 11(1), 105-134. https://doi.org/10.14421/al-athfal.2025.111-08

Laksana, D. A. W., Widodo, A. S., & Ardianto, D. T. (2024). Gambar imajinasi anak-anak sebagai aset dalam pengembangan animasi edukasi. *ANDHARUPA: Jurnal Desain*

Komunikasi Visual & Multimedia, 9(4), 488-502. https://doi.org/10.33633/andharupa.v9i4.9452

Laksana, S. D., et al. (2025). Enhancing reading literacy through augmented reality among Indonesian migrant children in Malaysia. *Journal of Integrative Elementary Education*, 5(2), 469-487. https://doi.org/10.21580/jieed.v5i2.26498

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, 66(5), 1141-1164. https://doi.org/10.1007/s11423-018-9581-2

Mayer, R. E. (2009). *Multimedia learning (2nd ed.).* Cambridge University Press.

Nieveen, N., & Folmer, E. (2013). *Formative evaluation in educational design research*. In Educational design research (pp. 152-169). SLO.

OECD. (2023). *PISA 2022 Results (Volume I and II)* - Country Note: Indonesia. OECD. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes

Richey, R. C., Klein, J. D., & Nelson, W. A. (2003). *Developmental research: Studies in instructional design and development.* Association for Educational Communications and Technology.

Rini, D. R., Wisesa, A. M., & T. M., R. D. (2024). Storytelling book for early childhood education: Recognize animal illustration. *KnE Social Sciences*. https://doi.org/10.18502/kss.v9i15.16182

Shofiyah, N., Jatmiko, B., Suprapto, N., Prahani, B. K., & Anggraeni, D. M. (2025). The use of technology to scientific reasoning in science education: A bibliometric and content analysis. *Social Sciences & Humanities Open*, 11, 101534. https://doi.org/10.1016/j.ssaho.2025.101534

Sugihyono, S. (2025). The dark side of AI in Indonesian education: Understanding its impact on the Alpha generation. *Literacy: International Scientific Journals of Social, Education and Humanities,* 4(1), 218-255. https://doi.org/10.56910/literacy.v4i1.2096

Supiani, S., Sumarno, S., Saputro, B. A., & Nurfitriah, N. (2025). Pengembangan media Wordwall berbasis gamifikasi untuk literasi permulaan anak usia 5-6 tahun. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini,* 9(6), 2507-2517. https://doi.org/10.31004/obsesi.v9i6.7498

Tarigan, F. N., Hasibuan, S. A., & Nurmayana, N. (2024). Application and challenges of digital storytelling based artificial intelligence for language skills: A narrative review. *SALTEL*, 7(1). https://doi.org/10.35307/saltel.v7i1.117

Tsang, S., Miao, R., Xiao, J., & Xiong, H. (2025). AnimAlte: Designing AI-infused cartoon videos to improve preschoolers' language learning. *arXiv*. https://doi.org/10.48550/arXiv.2502.12526

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636. https://doi.org/10.3758/BF03196322

Yanti, N., Arono, A., Hiasa, F., Junaidi, F., Noermanzah, N., & Kurniawan, R. (2025). Indonesian teachers' roles in designing and utilizing AI-powered animated videos. *International Journal of Learning Teaching and Educational Research*, 24(6), 833-856. https://doi.org/10.26803/ijlter.24.6.38