

The Effectiveness of a Knowledge Transfer Workshop on Motion-Augmented Reality Interactive Storybook Development: A Pre–Post Evaluation among Early Childhood Educators

Mariyati Mohd Nor, Ariza Abdullah, Nor Ili Ahmad Tajuddin, Nur Akmal Othman, Athira Najwa Zakaria, & Nur Syuhada Mat Sin

Universiti Pendidikan Sultan Idris (UPSI), Malaysia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.100900068>

Received: 16 September 2026; Accepted: 21 September 2026; Published: 30 September 2026

ABSTRACT

This study evaluated a Knowledge Transfer Programme (KTP) workshop designed to develop educators' capacity to conceptualise, design, and assemble Motion-Augmented Reality (MAR) interactive storybooks. The intervention combined generative artificial intelligence (GenAI) for story development and prompt-assisted scripting, AI-supported motion generation, and an augmented-reality authoring platform for linking printed pages with digital media. The evaluation focused on three competency domains: concept and narrative knowledge, pedagogical integration, and technical skill mastery. A single-group pre-test–post-test design was used with 19 participants comprising language teachers, preschool teachers, higher-education lecturers, and pre-service educators. Participants completed 5-point Likert-scale evaluation items before and after the workshop. Paired-samples t-tests were used to examine mean differences, with Wilcoxon signed-rank analysis identified as an appropriate non-parametric sensitivity analysis for paired observations. Across the three domains, post-workshop means were higher than pre-workshop means, with statistically significant differences reported for concept and narrative knowledge, pedagogical integration, and technical skill mastery (all $p < .001$). Effect sizes were large to very large (Cohen's $d = 1.43$ – 2.33). The largest domain-level gain was observed for technical skill mastery. Open-ended responses further indicated gains in cognitive structuring, practical production skills, and confidence, while also identifying constraints involving generative-AI usage limits, character consistency, and limited workshop time. The findings suggest that a scaffolded, hands-on professional learning model can support educators in moving from consuming digital materials to authoring contextually relevant MAR resources. However, because the study used a small, non-random single-group sample and self-reported competency measures, the findings should be interpreted as evidence of short-term programme-level change rather than causal evidence of classroom impact. The study contributes an applied model for combining AR, GenAI, and teacher professional development in early childhood and language education.

Keywords: augmented reality; motion-Augmented Reality; interactive storybooks; generative artificial intelligence; teacher professional development; knowledge transfer; early childhood education; digital literacy

INTRODUCTION

Background and Rationale

Digital technologies are increasingly being incorporated into early childhood learning environments, but their educational value depends on how they are designed and pedagogically integrated. Augmented reality (AR) is particularly relevant because it overlays digital information such as animation, audio, video, images, and three-dimensional objects onto the physical environment rather than requiring learners to abandon the physical learning context. A recent systematic review of AR and virtual-reality applications reported broad interest in immersive learning while also highlighting the importance of instructional design, educator capability, and contextual conditions in determining educational value (Sakr & Abdullah, 2024). A large meta-analysis of 134 quasi-experimental and experimental studies similarly found positive effects of AR across response,

knowledge/skill, and performance outcomes, although effects varied according to characteristics such as treatment duration and the affordances of the AR environment (Garzón et al., 2023).

These affordances are particularly relevant to storybook-mediated learning. In early childhood education, storybooks provide opportunities for language development, comprehension, vocabulary growth, narrative sequencing, and social-emotional learning. AR can extend the conventional storybook by connecting printed illustrations to audiovisual and animated content. In a study involving 90 preschool children, Şimşek (2024) compared printed and AR storybook experiences and examined story comprehension and retelling. The study provides direct evidence that AR storybooks can be investigated not simply as technological novelties but as learning resources linked to specific literacy outcomes. A systematic review focused specifically on preschool education likewise identified AR as an emerging didactic resource and emphasized the need to examine how AR is integrated into preschool teaching rather than treating the technology itself as the intervention (Hurtado-Mazeyra et al., 2024).

Despite this potential, creating high-quality interactive storybooks can be technically demanding. Educators may need to develop a narrative, write and adapt scripts, generate or edit visual assets, create motion, manage audio, prepare image targets, configure AR links, test the experience, and align the final resource with curriculum and learner needs. Consequently, the professional-development question is not merely whether teachers can use AR, but whether they can acquire sufficient technological, pedagogical, and content-related knowledge to author and integrate AR resources. This concern resonates with the Technological Pedagogical Content Knowledge (TPACK) perspective, in which effective technology integration depends on the interaction among technological knowledge, pedagogical knowledge, and content knowledge rather than on technical skill alone (Mishra & Koehler, 2006).

The emergence of generative AI has altered the feasibility of this authoring process. Large language models can support brainstorming, story structuring, drafting, revision, prompt development, and adaptation of materials, while multimodal generative systems can support image and video production. Recent reviews indicate that GenAI in education is being used for pedagogical enhancement, resource generation, professional learning, instructional design, and personalised support, although ethical, privacy, reliability, and professional-judgement issues remain important (Yusuf et al., 2024; UNESCO, 2023). A 2025 systematic review of GenAI in K–12 education also reported rapid growth in applications while noting the continuing need for pedagogically grounded implementation and stronger empirical evidence (Marzano, 2025).

For teachers, therefore, the central challenge is becoming a competent human author and pedagogical decision-maker who can use AI as an assistive technology. Recent research on teacher development suggests that GenAI-supported professional learning may strengthen teacher knowledge, instructional design, confidence, and AI literacy, but the evidence base remains relatively new and often involves small samples and short interventions (Tan, 2025). Similarly, a 2026 systematic review of teacher GenAI research reported that AI literacy, pedagogical knowledge, self-efficacy, institutional support, professional development, and teacher autonomy are important conditions shaping meaningful integration (Ayubian et al., 2026).

Against this background, Universiti Pendidikan Sultan Idris (UPSI) implemented a targeted KTP workshop to equip educators with a practical workflow for producing MAR interactive storybooks. The workshop combined story architecture, GenAI-assisted scripting, AI-supported motion generation, and AR marker-based deployment. The present evaluation examines whether this short, hands-on intervention was associated with measurable changes in participants' perceived competency across knowledge, pedagogical integration, and technical skill domains.

Problem Statement

The literature suggests that AR can support engagement and learning, yet there is a persistent implementation gap between the availability of immersive technologies and educators' ability to create and integrate meaningful resources. In preschool contexts, systematic evidence remains more limited than in higher education and other formal educational sectors (Hurtado-Mazeyra et al., 2024). This gap is important because early childhood educators often need resources that are developmentally appropriate, visually coherent, linguistically accessible, and closely connected to classroom routines.

Three practical problems motivated the KTP intervention. First, conventional AR production can involve multiple software environments and specialised skills, creating a barrier for educators without formal backgrounds in animation, programming, or digital media. Second, even when teachers can operate a tool, technical competence alone does not guarantee pedagogical integration; teachers must decide when, why, and how AR contributes to learning objectives. Third, the rapid availability of GenAI introduces both opportunities and new forms of complexity, including prompt quality, consistency of generated characters and scenes, intellectual-property considerations, privacy, output reliability, and dependence on commercial usage quotas. UNESCO (2023) therefore recommends a human-centred approach to GenAI that places pedagogical judgement, safety, privacy, equity, and ethical use at the centre of implementation.

Accordingly, there is a need for programme-level evidence showing whether a structured, scaffolded workshop can produce measurable short-term changes in educators' confidence and perceived competence in MAR storybook development. The present study addresses this need by evaluating a single KTP workshop using pre- and post-programme measures and qualitative feedback.

Research Objectives

RO1: To determine changes in participants' concept and narrative knowledge related to MAR storybook development, including MAR concepts, visual composition, script scoring, narrative structuring, character profiling, and CEFR-oriented difficulty alignment.

RO2: To examine changes in participants' perceived confidence and pedagogical readiness to integrate MAR storybooks into lesson planning and differentiated language instruction.

RO3: To determine changes in participants' perceived technical skill in using GenAI-assisted storyboarding, AI-based motion generation, camera and sound controls, AR marker configuration, and physical–digital storybook assembly.

RO4: To identify participants' perceived cognitive, psychomotor, and affective learning gains and the operational constraints encountered during the workshop.

LITERATURE REVIEW

Augmented Reality and Learning

AR is generally distinguished from virtual reality because digital information is superimposed on, or linked to, the user's physical environment. This characteristic makes AR particularly compatible with physical learning materials such as books, worksheets, laboratory objects, and classroom displays. The educational value of AR lies not in immersion alone but in its capacity to connect representations, actions, and contextual information in ways that may support attention, interaction, visualisation, and learning (Akçayır & Akçayır, 2017; Garzón et al., 2023).

Evidence from systematic reviews supports a generally positive but non-uniform relationship between AR and learning. Garzón et al. (2023), synthesising 134 experimental and quasi-experimental studies, reported benefits across affective/response, knowledge and skill, and performance outcomes. Sakr and Abdullah (2024) similarly concluded that AR and related immersive technologies can create new opportunities for learners and educators, while stressing that implementation quality and educational context matter. More recent review evidence also indicates that AR can enhance engagement, motivation, and interactive learning, while effects on achievement may vary by subject, instructional function, and design (Dhaas, 2024).

The implications for the present study are twofold. First, the workshop should not be evaluated solely by whether participants can operate an AR platform. Second, the professional-development intervention should develop a chain of competencies extending from conceptual understanding to pedagogical planning and technical production. This logic underpins the three-domain structure used in the present evaluation.

AR Storybooks in Early Childhood and Literacy

Storybooks provide a natural context for AR because the physical page can function as the anchor for additional digital content. AR can add movement, narration, sound effects, character dialogue, and visual cues without

eliminating the tactile and social affordances of print reading. The study by Şimşek (2024) is especially relevant because it examined preschool children's comprehension and retelling following exposure to printed versus AR storybooks. Such evidence supports the use of storybooks as a meaningful context for evaluating AR, while also suggesting that the quality of the story and the alignment between digital overlays and narrative objectives remain critical.

AR storybooks have also been explored for purposes beyond conventional literacy. Sari et al. (2022) developed an AR storybook for early childhood financial literacy, illustrating how contextual narratives can be combined with interactive digital representations to make abstract concepts more accessible. Lin et al. (2023) examined home-based AR storybook training in relation to emotional functioning and socialisation among children with autism spectrum disorder. Collectively, these studies indicate that AR storybooks can support cognitive, language, and socio-emotional purposes, although the specific outcomes depend on the pedagogical design and learner population.

A key implication is that teachers need authoring competence, not merely consumer competence. If teachers can design their own storybooks, they can potentially adapt language complexity, cultural references, characters, pacing, and learning tasks to local curricular requirements. In the present workshop, this principle was operationalised through CEFR A1/A2-oriented narrative design and the integration of language-learning considerations into story development.

Generative AI as an Authoring and Professional-Learning Tool

GenAI has expanded the range of content that educators can produce without advanced programming or graphic-design expertise. Large language models can assist with ideation, story outlines, dialogue, character profiles, and prompt refinement; multimodal systems can support visual and motion asset generation. A systematic mapping review of 407 publications found that GenAI research in education has concentrated on applications and impacts, ethical issues, adoption, performance, writing support, pedagogical enhancement, and professional skills development (Yusuf et al., 2024).

At the same time, GenAI should not be treated as an autonomous instructional designer. UNESCO (2023) advocates a human-centred approach because generative systems can produce inaccurate, biased, privacy-sensitive, or otherwise unsuitable outputs. In teacher professional development, this means educators need not only operational skills but also the capacity to evaluate, revise, contextualise, and ethically deploy AI-generated content. Recent reviews of AI-focused teacher professional development similarly conclude that technical training alone is insufficient; pedagogical knowledge, AI literacy, institutional support, and continuing professional learning are also necessary (Tan, 2025; Marzano, 2025).

This perspective also helps explain the workshop's emphasis on prompt engineering, character consistency, narrative structure, and curriculum alignment. These activities position GenAI as a tool within an educator-led design process rather than as a replacement for professional judgement.

Teacher Professional Development, Self-Efficacy and TPACK

Teacher professional development is more likely to influence classroom practice when it combines relevant content, active participation, opportunities for practice, feedback, and contextual application. Technology-enabled professional development can support active learning and collaboration, although evidence also highlights the importance of expert support, feedback, reflection, equity, and sustained implementation (Huang et al., 2024). A meta-analysis of professional development for in-service STEM teachers found a positive overall effect on teacher self-efficacy, illustrating the potential of structured professional learning to influence teachers' beliefs about their capability to teach with confidence (Zhou et al., 2023).

TPACK provides a useful conceptual lens for interpreting the three competency domains in this study. Mishra and Koehler (2006) argue that meaningful technology integration involves the interaction of technology, pedagogy, and content rather than isolated technological knowledge. The present workshop extends this logic to MAR storybook production: participants first need conceptual and narrative knowledge, then pedagogical knowledge about integrating the resource into lessons, and finally technical competence to produce and deploy

the resource.

The three domains should therefore be interpreted as related rather than independent. Technical mastery without pedagogical integration may produce visually attractive but instructionally weak materials; pedagogical knowledge without technical capacity may limit teachers' ability to author the materials they envision. The workshop was designed to reduce this gap through a sequential, hands-on workflow.

Research Gap

Three gaps emerge from the literature reviewed above. First, much AR research evaluates learner outcomes after an intervention, whereas fewer studies examine the professional learning required for educators to become AR content authors. Second, research on GenAI in education is expanding rapidly, but empirical evidence on short, applied teacher-development interventions remains relatively limited and often involves small samples (Tan, 2025). Third, the intersection of GenAI-assisted authoring, AR storybooks, and early childhood/language education remains underexplored. The present study contributes programme-level evidence by examining an applied KTP model that integrates all three components.

METHODOLOGY

Research Design

The study employed a single-group pre-test–post-test evaluation design. Nineteen participants completed programme evaluation measures before and after the KTP workshop. The design was selected to assess short-term change in perceived competency following an intensive, hands-on professional-learning experience. Because the same participants were assessed at two time points, the primary comparison concerns within-person change rather than differences between independent groups.

The original manuscript described both a paired-samples t-test and a Mann–Whitney U test. Because the observations are paired pre- and post-workshop responses from the same participants, the appropriate non-parametric counterpart is the Wilcoxon signed-rank test rather than the Mann–Whitney U test, which is designed for independent groups. The revised manuscript therefore reports the paired-samples t-test as the parametric analysis and treats the Wilcoxon signed-rank test as the appropriate non-parametric sensitivity analysis where required. No new Wilcoxon statistics are fabricated here because the underlying paired item-level data were not provided in the manuscript.

The design does not include a control group, random assignment, or follow-up measurement. Accordingly, statistically significant pre–post differences should be interpreted as evidence of change during the programme period, not as definitive causal evidence that the workshop alone produced the observed changes.

Participants

The workshop cohort consisted of 19 participants from four professional categories: language teachers ($n = 6$), preschool teachers ($n = 5$), higher-education lecturers ($n = 5$), and pre-service trainees ($n = 3$). Eight participants held a master's degree, five held a doctorate, four held a bachelor's degree, and two held a diploma. Ten participants reported more than 10 years of teaching experience, three reported 7–10 years, one reported 1–3 years, and five were either students or had less than one year of experience. Regarding prior AR/animation experience, eight participants reported no prior experience, seven reported basic knowledge, and four reported intermediate skill.

Table 1: Participant

Characteristic	Category	n (%)
Professional role	Language teachers	6 (31.6%)
	Preschool teachers	5 (26.3%)

	Higher-education lecturers	5 (26.3%)
	Pre-service trainees	3 (15.8%)
Education	Master's degree	8 (42.1%)
	Doctorate	5 (26.3%)
	Bachelor's degree	4 (21.1%)
	Diploma	2 (10.5%)
Teaching experience	>10 years	10 (52.6%)
	7–10 years	3 (15.8%)
	1–3 years	1 (5.3%)
	<1 year/student	5 (26.3%)
Prior AR/animation	No prior experience	8 (42.1%)
	Basic knowledge	7 (36.8%)
	Intermediate	4 (21.1%)

Workshop Architecture

The workshop followed a sequential production pipeline consisting of three modules.

Module 1 – Storytelling architecture and AI-assisted scripting: Participants developed story outlines, character profiles, scripts, prompts, and visual composition plans. The activities included consideration of CEFR A1/A2 language difficulty, narrative sequencing, loglines, plot development, and visual framing.

Module 2 – Motion generation: Participants transformed scene descriptions into motion assets using an AI-based generation workflow. Attention was given to prompt structure, camera movement, character actions, timing, and audio-related considerations.

Module 3 – AR marker layering and physical assembly: Participants connected printed storybook pages with digital assets using Artivive, tested target recognition and playback, and assembled the physical–digital storybook.

The sequence was intentionally scaffolded from conceptualisation to production. This reflects the principle that professional learning should provide opportunities to learn concepts, practise procedures, and apply knowledge to an authentic product. It also operationalises the TPACK logic described above by requiring participants to connect content and narrative decisions with pedagogical purposes and technical execution.

Instrument and Measures

The evaluation instrument used a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The manuscript reports 14 competency items distributed across three domains: concept and narrative knowledge (five items), pedagogical integration (four items), and technical skill/AI–AR tools (five items). Post-programme evaluation also included 10 workshop-experience indicators covering clarity of objectives, content alignment, professional relevance, knowledge acquisition, technical competence, confidence, presentation quality, perceived impact on attention, usefulness for resource shortages, and self-efficacy in autonomous AI/AR tool execution.

To ensure content validity, the evaluation instrument was adapted and adopted from established technology integration and competency frameworks. The initial draft was subjected to an expert review process involving five subject-matter experts (SMEs) in early childhood education, educational technology, and language pedagogy. All five experts evaluated the relevance, clarity, and suitability of the adapted items, reaching full consensus (100% agreement) on all 14 competency items prior to administration. While content validity was explicitly established through expert judgment, formal psychometric properties such as construct validity via factor analysis and internal consistency reliability (e.g., Cronbach's alpha) were not executed in this baseline evaluation due to the small sample size ($n = 19$). Future studies utilizing larger participant samples will incorporate rigorous psychometric validation procedures to confirm the structural properties of the instrument.

Data Analysis

Descriptive statistics were used to summarise pre- and post-workshop means and standard deviations. Paired-samples t-tests were used to evaluate mean differences across the reported competency domains, and Cohen's d was used to quantify standardised change. The manuscript reports domain-level effect sizes between 1.43 and 2.33. Given the small sample and ordinal response format, a non-parametric paired analysis using the Wilcoxon signed-rank test is also appropriate as a robustness check. However, the raw paired observations were not available in the manuscript supplied for revision, so corresponding Wilcoxon statistics are not reported.

Interpretation focuses on the direction and magnitude of change rather than statistical significance alone. This is consistent with current recommendations for educational intervention research, where effect size and contextual relevance provide information that p-values alone cannot convey. Because multiple domain and item-level comparisons were conducted, the possibility of inflated Type I error should also be acknowledged.

RESULTS

Domain-Level Competency Change

All four reported competency outcomes showed higher post-workshop means than pre-workshop means. The largest domain-level change was observed for Technical Skill Mastery, followed by Concept and Narrative Knowledge and Pedagogical Integration. The overall educational competency score also increased substantially.

Table 2 : Domain-Level Competency Change

Domain	Pre M (SD)	Post M (SD)	t	p	Cohen's d
Concept & Narrative Knowledge	2.78 (1.05)	4.45 (0.43)	-6.46	< .001	2.10
Pedagogical Integration	3.29 (0.92)	4.41 (0.61)	-4.39	< .001	1.43
Technical Skill Mastery	2.41 (1.05)	4.32 (0.49)	-7.19	< .001	2.33
Overall Educational Competency	2.79 (0.88)	4.39 (0.46)	-7.08	< .001	2.30

The Concept and Narrative Knowledge domain increased by 1.67 scale points, from 2.78 to 4.45, with a reported large standardised effect ($d = 2.10$). Pedagogical Integration increased by 1.12 points, from 3.29 to 4.41 ($d = 1.43$). Technical Skill Mastery increased by 1.91 points, from 2.41 to 4.32, producing the largest reported effect ($d = 2.33$). These results indicate that the workshop was associated with substantial short-term improvement in participants' perceived capability across conceptual, pedagogical, and technical dimensions.

Item-Level Chang

The item-level results show positive change across all 14 reported competency items. The largest gains were reported for physical-digital hybrid storybook compilation (+2.16), AI-based motion command and animation generation (+2.15), character profiling and Rule-of-Thirds framing (+2.11), and Artivive marker configuration and asset uploading (+2.10).

Table 3 : Item – level Change

Domain / Item	Pre M (SD)	Post M (SD)	Gain
1.1 MAR definition and core concepts	2.79 (1.44)	4.58 (0.51)	+1.79
1.2 Visual (active/static), audio & AR comprehension	2.84 (1.38)	4.47 (0.51)	+1.63
1.3 Script scoring, logline & plot-twist design	2.95 (0.97)	4.47 (0.51)	+1.52
1.4 CEFR difficulty alignment (A1/A2)	3.05 (1.18)	4.37 (0.60)	+1.32
1.5 Character profiling & Rule-of-Thirds framing	2.26 (1.28)	4.37 (0.50)	+2.11
2.1 MAR integration into RPH	3.22 (1.22)	4.37 (0.68)	+1.15
2.2 Managing active tech-supported classroom space	3.26 (1.10)	4.32 (0.75)	+1.06
2.3 Differentiated instruction based on comprehension	3.26 (0.81)	4.47 (0.61)	+1.21
2.4 Integrating listening, speaking, reading & writing	3.37 (0.90)	4.47 (0.51)	+1.10
3.1 AI collaborative storyboarding	3.16 (1.01)	4.42 (0.51)	+1.26
3.2 AI motion commands & animation generation	2.32 (1.29)	4.47 (0.61)	+2.15
3.3 Camera control & sound engineering	2.26 (1.15)	4.11 (0.66)	+1.85
3.4 Artivive marker configuration & asset uploading	2.16 (1.17)	4.26 (0.73)	+2.10
3.5 Physical–digital hybrid storybook compilation	2.16 (1.21)	4.32 (0.67)	+2.16

Programme Satisfaction and Perceived Utility

Participants also reported high levels of satisfaction with the workshop. The mean ratings for the 10 reported experience indicators ranged from 4.58 to 4.74. The highest ratings concerned alignment between workshop content and objectives, relevance to professional pedagogical needs, acquisition of theoretical knowledge, acquisition of hands-on technical competence, and clarity of presentation. Confidence in future MAR use and self-efficacy in autonomous AI/AR execution were also rated positively.

Table 4 : Programme satisfaction

Indicator	M (SD)
Objectives clearly communicated	4.68 (0.48)
Content aligned with objectives	4.74 (0.45)
Content matched professional needs	4.74 (0.45)
Substantive new theoretical knowledge acquired	4.74 (0.45)
Practical technical competencies acquired	4.74 (0.45)
Confidence using MAR in future teaching	4.68 (0.48)

Presentation was clear and systematic	4.74 (0.45)
Perceived impact on student attention/focus	4.63 (0.50)
Usefulness for instructional-resource shortages	4.58 (0.61)
Self-efficacy for autonomous AI/AR execution	4.58 (0.61)

Qualitative Findings

Open-ended post-workshop responses were organised in the original evaluation around two overarching themes: tri-domain learning impact and operational bottlenecks. The qualitative findings complement the quantitative results by illustrating what participants perceived to have changed and what prevented a smoother production process.

Cognitive learning

Participants described a shift from viewing stories as linear text toward structuring stories as multimodal sequences involving plot development, visual composition, movement, sound, and character action. Prompt engineering was perceived as useful for translating narrative intentions into more specific image and motion-generation instructions. This finding is consistent with the broader literature suggesting that GenAI can support idea generation, content development, and instructional-resource production, while still requiring human evaluation and revision (Yusuf et al., 2024; UNESCO, 2023).

Psychomotor and technical learning

Participants reported hands-on gains in navigating AI and AR interfaces, generating motion assets, configuring markers, uploading digital resources, and aligning physical pages with digital overlays. These observations are particularly relevant because the largest quantitative domain gain occurred in technical skill mastery. The combination of demonstration, guided practice, and immediate production appears to have provided an authentic context for learning the technical workflow.

Affective learning and confidence

Participants also reported increased confidence, motivation, and enthusiasm for creating digital learning resources. Such affective changes are relevant to technology integration because teachers' beliefs about their capability are associated with technology-integration knowledge and practice. Evidence from meta-analytic research indicates a positive relationship between teachers' technology-integration self-efficacy and TPACK (Liu et al., 2022), while professional-development interventions can improve teacher self-efficacy (Zhou et al., 2023).

Operational constraints

- **Usage limits and subscription costs:** Participants reported encountering quotas or restrictions associated with free-tier AI services when producing multiple assets. This suggests that institutional access arrangements may become a practical determinant of workshop continuity.
- **Character and visual consistency:** Maintaining the same appearance and characteristics of a person or fictional character across sequential AI-generated scenes was reported as challenging. This is a recognised practical issue in generative visual workflows and can affect narrative coherence.
- **Limited production time:** Prompt iteration, animation refinement, audio adjustment, and AR testing required more time than a single short workshop could comfortably provide. This suggests that future KTP delivery should distinguish introductory orientation from an extended production phase.

DISCUSSION

The Workshop as a Scaffolded Professional-Learning Intervention

The central finding is that participants reported substantially higher competency scores after the workshop across conceptual, pedagogical, and technical domains. The result is consistent with the broader literature on technology-enabled professional development, which indicates that structured learning opportunities can strengthen teachers' knowledge and skills when they include active learning, collaboration, expert support, and opportunities for application (Huang et al., 2024). The findings also align with evidence that professional development can improve teachers' self-efficacy (Zhou et al., 2023).

The workshop's design may help explain the pattern of results. Rather than presenting AR as an isolated tool, the intervention required participants to complete a coherent production cycle: conceptualise a story, formulate prompts, generate assets, integrate movement and audio, configure AR markers, and test the final product. This type of authentic task can make technology learning immediately relevant because each technical action is connected to a tangible educational product.

The results should nevertheless be interpreted cautiously. Because the evaluation measured perceived competency using self-report items, a high post-test mean does not necessarily indicate equivalent mastery in an objective performance task. Future studies should therefore combine self-report scales with rubric-based assessment of completed storybooks, observation of independent tool execution, and delayed follow-up measures.

Interpreting the Technical Skill Gain

Technical Skill Mastery recorded the largest domain-level change (1.91 points; $d = 2.33$). This result is notable because 42.1% of participants reported no prior AR/animation experience. A plausible interpretation is that participants benefited from explicit modelling and immediate guided practice. However, the study design does not allow the magnitude of the change to be attributed solely to the workshop, because maturation, familiarity with the instrument, peer support, and other contextual influences cannot be separated from the intervention.

The result nevertheless has practical significance for KTP design. If educators with limited prior experience can report substantial short-term gains after a structured workshop, an institutional training model may be feasible. The next step is to determine whether these gains transfer to independent production and actual classroom use. This is particularly important because reviews of teacher AI integration emphasise that access and short-term confidence are not sufficient; sustained professional development, institutional support, autonomy, and pedagogical knowledge also shape implementation (Ayubian et al., 2026).

Pedagogical Integration as the Critical Bridge

Pedagogical Integration increased from 3.29 to 4.41 ($d = 1.43$), a substantial change but smaller than the technical domain. This pattern is theoretically meaningful. Producing an AR artefact and integrating it into a lesson are related but different competencies. The latter requires decisions about learning objectives, timing, classroom management, differentiation, learner interaction, and assessment. TPACK provides a useful explanation because effective technology integration emerges from the intersection of technological, pedagogical, and content knowledge (Mishra & Koehler, 2006).

The relatively high post-workshop scores should therefore be interpreted as increased perceived readiness rather than evidence that participants had already developed stable classroom implementation expertise. Future KTP cycles should include lesson-plan development, microteaching, peer review, and classroom trials so that pedagogical integration can be observed rather than inferred from self-report.

GenAI as an Enabler, not a Replacement for Teacher Expertise

The workshop illustrates a potentially important shift in digital-content production: educators can use GenAI to reduce some of the technical and time barriers associated with drafting and prototyping content. This is consistent

with recent systematic reviews identifying resource development, lesson design, professional learning, and personalised support as prominent educational uses of GenAI (Yusuf et al., 2024; Tan, 2025).

However, the workshop's operational challenges also demonstrate why AI literacy must include critical evaluation. Generated scripts, images, and videos may require revision for linguistic accuracy, cultural appropriateness, developmental suitability, representation, and consistency. UNESCO (2023) specifically emphasises human-centred, safe, equitable, and meaningful use of GenAI. Thus, the desired outcome of the KTP is not educator dependence on AI, but educator capacity to direct, evaluate, modify, and integrate AI outputs responsibly.

Contribution to Early Childhood and Language Education

The study contributes an applied example at the intersection of early childhood education, language education, AR, and GenAI. Its practical contribution is the articulation of a relatively accessible authoring pipeline that begins with narrative and pedagogical decisions and ends with a physical–digital learning resource. This is particularly relevant where educators need locally contextualised materials that may not be readily available commercially.

The approach also creates possibilities for multilingual and CEFR-oriented content. Story complexity, vocabulary, dialogue length, pronunciation support, and visual scaffolding can be adapted to language level. Nevertheless, the current study evaluates educator competency rather than children's actual language outcomes. Claims about improved child learning should therefore remain prospective until the resulting storybooks are tested with appropriate child participants and outcome measures.

IMPLICATIONS AND RECOMMENDATIONS

Implications for KTP and Teacher Professional Development

- Adopt a two-stage model: an introductory conceptual workshop followed by an extended production clinic in which participants complete and refine one MAR storybook.
- Use a TPACK-informed competency framework that explicitly assesses technology, pedagogy, content, and the intersections among them.
- Include objective performance assessment using a rubric covering narrative quality, language appropriateness, visual coherence, technical functionality, AR usability, and pedagogical alignment.
- Provide structured peer feedback and microteaching opportunities so participants can test whether the storybook functions effectively as a learning activity.
- Include AI literacy and responsible-use components covering privacy, verification of generated content, copyright/intellectual-property awareness, bias, cultural representation, and appropriate human oversight, consistent with UNESCO guidance.
- Where feasible, provide institutional access to required software or establish a common platform environment to reduce disruption from individual account quotas and subscription constraints.
- Develop reusable templates for prompts, storyboards, character profiles, lesson plans, AR marker layouts, and assessment rubrics so participants can focus on pedagogy rather than repeatedly solving technical problems.

RECOMMENDATIONS FOR FUTURE RESEARCH

To address the methodological constraints of this preliminary study and establish stronger evidence of programme efficacy, future research should expand upon several key dimensions:

1. **Sample Scale and Diversity:** Replicate the workshop across larger and more diverse educator cohorts spanning multiple institutions, geographic regions, and varied educational settings (e.g., public vs. private preschools, rural vs. urban centers).

2. **Controlled Research Designs:** Move beyond the single-group pre–post evaluation toward quasi-experimental or randomized controlled trial (RCT) designs with appropriate comparison groups to isolate the intervention's true effect from confounding variables.
3. **Objective and Multi-Source Assessments:** Supplement self-report competency surveys with objective performance metrics, including rubric-based evaluations of completed MAR storybooks, direct observational assessments during independent tool usage, and classroom implementation artifacts.
4. **Longitudinal Assessment:** Implement delayed post-test follow-ups (e.g., at 3 or 6 months post-workshop) to measure the retention and sustainability of teacher competency gains over time.
5. **Direct Child-Level Outcomes:** Investigate how educators' MAR authoring competence translates into actual learner outcomes, measuring young children's language acquisition, story comprehension, classroom engagement, and early digital literacy during classroom implementation.

CONCLUSION

This evaluation indicates that a structured, hands-on KTP workshop was associated with substantial short-term increases in participants' perceived concept knowledge, pedagogical integration, and technical skill in developing MAR interactive storybooks. Technical skill showed the largest reported gain, while qualitative feedback indicated complementary cognitive, psychomotor, and affective development. At the same time, the findings identified practical barriers involving platform quotas, subscription costs, visual consistency, and limited production time.

The principal contribution of the intervention is not simply the introduction of AR or GenAI as new technologies. Rather, it demonstrates a professional-learning approach in which educators are guided through a complete authoring workflow and encouraged to connect technological production with narrative, language, and pedagogical decisions. This approach is consistent with TPACK-oriented technology integration and with recent evidence emphasising that effective teacher AI adoption requires more than access to tools.

Because the study used a small single-group pre–post design and self-reported measures, the results should be interpreted as promising programme-evaluation evidence rather than definitive causal evidence of classroom effectiveness. Future research should strengthen the design through comparison groups, objective performance assessment, follow-up measurement, and direct evaluation of children's learning outcomes. Such work would help establish whether educator competency gains translate into sustainable classroom practice and meaningful early childhood learning outcomes.

ACKNOWLEDGMENT

This manuscript is based on the research entitled “MotionAR Interactive Storybook-Jejak Malaysia dalam Bahan Penceritaan Bahasa Arab” The researches would like to thank the Faculty Of Language and Communication and The Research Management and Innovation Centre (RMIC), Universiti Pendidikan Sultan Idris (UPSI), Research RIMC (Code: 2026-0029-106-01-1) which funded this research.

REFERENCES

1. Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1–11. <https://doi.org/10.1016/j.edurev.2016.11.002>
2. Ayubian, S., Aguilar Chávez, G., Wiafe, E., Shahsavari, V., & Clark, J. S. (2026). Teachers and generative AI: A systematic review of adoption, competence, professional development, and socio-technical experiences. *Education Sciences*, 16(9), 1501. <https://doi.org/10.3390/educsci16091501>
3. Dhaas, A. (2024). Augmented reality in education: A review of learning outcomes and pedagogical implications. *American Journal of Computing and Engineering*, 7(3), 1–18. <https://doi.org/10.47672/ajce.2028>
4. Garzón, J., Kinshuk, Baldiris, S., Gutiérrez, J., & Pavón, J. (2023). How do augmented reality

- applications affect students' academic achievement? A meta-analysis and meta-regression of 134 quasi-experimental and experimental studies. *Educational Research Review*, 39, 100513.
5. Huang, L., Liang, M., Xiong, Y., Wu, X., & Lim, C. P. (2024). A systematic review of technology-enabled teacher professional development during COVID-19 pandemic. *Computers & Education*, 223, 105168. <https://doi.org/10.1016/j.compedu.2024.105168>
6. Hurtado-Mazeyra, A., Condori-Yucra, N., Ponce-Alvarez, E., Limaymanta, C. H., & Suárez-Guerrero, C. (2024). Didactic use of augmented reality in preschool education: A systematic review. *Revista Complutense de Educación*. <https://doi.org/10.5209/rced.85815>
7. Lin, L.-Y., Lin, C.-H., Chuang, T.-Y., Loh, S. C., & Chu, S. Y. (2023). Using home-based augmented reality storybook training modules for facilitating emotional functioning and socialization of children with autism spectrum disorder. *International Journal of Developmental Disabilities*, 71(1), 87–94. <https://doi.org/10.1080/20473869.2023.2202454>
8. Liu, X., et al. (2022). The relationship between teachers' information technology integration self-efficacy and TPACK: A meta-analysis. *Educational and Information Technologies / related indexed source*.
9. Marzano, D. (2025). Generative Artificial Intelligence (GAI) in teaching and learning processes at the K-12 level: A systematic review. *Technology, Knowledge and Learning*, 31, 789–829. <https://doi.org/10.1007/s10758-025-09853-7>
10. Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
11. Qian, Y. (2025). Pedagogical applications of generative AI in higher education: A systematic review of the field. *TechTrends*, 69, 1105–1120. <https://doi.org/10.1007/s11528-025-01100-1>
12. Sakr, A., & Abdullah, T. (2024). Virtual, augmented reality and learning analytics impact on learners, and educators: A systematic review. *Education and Information Technologies*, 29, 19913–19962. <https://doi.org/10.1007/s10639-024-12602-5>
13. Sari, R. C., Aisyah, M. N., Ilyana, S., & Hermawan, H. D. (2022). Developing a financial literacy storybook for early childhood in an augmented reality context. *Contemporary Educational Technology*, 14(2), ep363. <https://doi.org/10.30935/cedtech/11734>
14. Şimşek, E. E. (2024). The effect of augmented reality storybooks on the story comprehension and retelling of preschool children. *Frontiers in Psychology*, 15, 1459264. <https://doi.org/10.3389/fpsyg.2024.1459264>
15. Tan, Q. (2025). Reimagining teacher development in the era of generative AI: A scoping review. *Teaching and Teacher Education*, 168, 105236. <https://doi.org/10.1016/j.tate.2025.105236>
16. UNESCO. (2023). Guidance for generative AI in education and research. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
17. Yusuf, A., et al. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, 12(2), e3489. <https://doi.org/10.1002/rev3.3489>
18. Zhou, X., Shu, L., Xu, Z., & Padrón, Y. (2023). The effect of professional development on in-service STEM teachers' self-efficacy: A meta-analysis of experimental studies. *International Journal of STEM Education*, 10, 37. <https://doi.org/10.1186/s40594-023-00422-x>