International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

Development Of The Ar-Visumathia Anaglyph 3d Adventure Model For The Stimulation Of Early Childhood Visual-Spatial And Mathematical Abilities

Desi Nurillah
Master of Early Childhood Education, Faculty of Education, Jakarta State University, Indonesia
Ade Dwi Utami
Master of Early Childhood Education, Faculty of Education, Jakarta State University, Indonesia
Sri Indah Pujiastuti
Master of Early Childhood Education, Faculty of Education, Jakarta State University, Indonesia
Erif Ahdhianto
Master of Early Childhood Education, Faculty of Education, Jakarta State University, Indonesia
Nurhamdayani
Master of Early Childhood Education, Faculty of Education, Jakarta State University, Indonesia
Rachel Carmenita
Early Childhood Education, Faculty of Education, Jakarta State University, Indonesia

Corresponding Author: desi.nuril@unj.ac.id

Abstract

This study examined the feasibility and classroom applicability of AR-VisuMathia, a web-based anaglyph 3D adventure model that a teacher guides on one shared classroom screen to stimulate visual-spatial and early mathematics skills in children aged 5-8. The article reports the expert validation and trial stages of a prototyping research and development study. One media expert and one content expert validated the product. A small-group trial involved 25 children, and a large-group trial involved 100 children in Bekasi, Indonesia. Data were analysed descriptively and with a paired t-test. Both experts rated the medium as highly feasible (media 95%, content 92%). Children's responses were highly positive in the small-group (91.6%) and large-group (95.0%) trials, but scores were so uniform that the questionnaire's internal consistency could not be demonstrated. Visual-spatial and early mathematics scores rose from 68.05 to 79.26, $t(99) = 6.96$, $p < .001$, $d = 0.70$. Without a comparison group, this gain cannot be attributed to the medium. The findings support the medium's classroom applicability and inform a controlled effectiveness study.

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

Keywords: Anaglyph 3D; Early Childhood; Early Mathematics; Learning Media; Visual-Spatial Skills

INTRODUCTION

Mathematical skills when children enter school are associated with mathematical achievement years later (Duncan et al., 2007), and planned math interventions from an early age have been shown to be beneficial (Clements & Sarama, 2011). Visual-spatial skills, such as understanding the position, direction, and arrangement of objects, can also be trained; meta-analyses of 217 studies showed the effects of spatial exercise that persisted and could move to other tasks (Uttal et al., 2013).

These two domains are interrelated. Preschoolers' ability to construct blocks is related to their initial mathematical ability (Verdine et al., 2014). Spatial-oriented geometry learning for 32 weeks improved children's spatial and numerical skills (Hawes et al., 2017), and visuospatial exercises improved the mathematical performance of elementary school students (Lowrie et al., 2017). A meta-analysis of 29 controlled studies confirmed that spatial exercises can improve math performance, with greater effects on older participants and on exercises using concrete objects (Hawes et al., 2022). Therefore, activities that combine spatial and numerical questions in one meaningful context are worth developing.

For Indonesia, strengthening this foundation is urgent. In PISA 2025, the average mathematics score of Indonesian students aged 15 years is 364, and only 17% of students achieve the minimum proficiency level, compared to 65% of the OECD average (OECD, 2026). This data concerns adolescents, so it does not show the cause at the early childhood level, but is a reason to strengthen the foundation from the beginning. In elementary school, students' numeracy literacy skills are also still a problem (Rakhmawati & Mustadi, 2022). The applicable curriculum places the basics of mathematics as part of the learning outcomes of the Foundation Phase (Ministry of Education, Culture, Research, and Technology, 2024).

AR-VisuMathia was developed to answer this need. The mode studied is a web-based 3D anaglyph adventure: a scene is seen with red-cyan glasses on one classroom screen, and the teacher guides the child's turn and discussion. In Indonesia, 3D media for spatial skills is widely developed in the form of *augmented reality applications*, such as a space-building application for spatial intelligence (Insani & Firdaus, 2024). AR-VisuMathia is different because it uses a single shared screen guided by the teacher. This form was chosen because interactive boards and classroom screens can support shared learning, although the impact is highly dependent on the

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

way the teacher guides (Higgins et al., 2007; Kennewell & Beauchamp, 2007). On the other hand, stereoscopic displays can cause visual discomfort in some viewers (Lambooij et al., 2009; McIntire et al., 2014; Read & Bohr, 2014), and evidence of the use of anaglyphs in education mostly comes from adult participants (Peker et al., 2022). Therefore, the feasibility and applicability of this media to children need to be tested, not assumed.

The previous article reported the analysis of teacher needs and the design of the AR-VisuMathia prototype (Nurillah et al., 2026). This article continues with the validation and trial stages. The research questions are: (1) how is the feasibility of media according to media experts and subject matter experts; (2) how do children respond to media in small group and large group trials; and (3) how do children's initial visual-spatial and mathematical ability scores change before and after using media in large group trials.

METHOD

This study uses a research and development design with a *seven-stage prototyping* model: needs analysis, initial design, prototyping, expert validation, trial, revision, and effectiveness testing. Stages 1-3 have been previously reported. This article reports stages 4-6 in the order of implementation in Table 1.

Table 1. Chronology of validation, revision, and trial implementation

Date	**Activities**
July 31, 2026	Media expert validation
August 10, 2026	Subject matter expert validation
August-October 2026	Product revision based on expert advice
October 29, 2026	Small group trial (25 children)
30 October-4 November 2026	Product revision based on small group tests
November 5-12, 2026	Large group trial (100 children)

The AR-VisuMathia model consists of three components: a web-based 3D anaglyph adventure app, a teacher-guided classroom flow, and a teacher's guide. The anaglyph mode app runs in a browser and is displayed on an interactive panel, Smart TV, or laptop connected to the screen. The teacher opens the app with the access key, selects the world, and then guides the child in turn answering on the screen while leading the discussion (Figure 1). The child wears red-

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

cyan glasses. The Vimo and Vima guides deliver instructions orally. One session is designed to last 15-20 minutes.

Figure 1. Web version of AR-VisuMathia: (a) the teacher access key screen and (b) the world select screen



Source: researcher documentation

Validation was carried out by one media expert (lecturer at the University of Education Indonesia) and one material expert (lecturer at the University of Muhammadiyah Prof. Dr. HAMKA, in the field of Early Childhood Education). The small group trial involved 25 children and the large group trial involved 100 children from partner schools in Bekasi City, aged 5-7 years. A total of 49 children in the large group test filled out a response questionnaire. The research was carried out in partner schools that have collaborated with the team since the needs analysis stage.

The research instruments are summarized in Table 2. The response questionnaire was adapted from an interactive media questionnaire for students. Each statement was read by the classroom teacher; children who could read short sentences also read it themselves. In the small group trial, some children were accompanied by researchers when filling out questionnaires. The same test was given before and after the child used the media.

Table 2. Research instruments

Instruments	Contents	Scale
Media expert validation sheet	20 items; aspects of display, media engineering, audio-visual communication, usefulness	1-4

Copyright at authors some right reserved this work is licensed under a Creative Common s Attribution-Share Alike 4.0 International License

Instruments	Contents	Scale
Subject matter expert validation sheet	25 items in 6 aspects, plus a review of 12 scenes	1-4
Children's response questionnaire	20 items; aspects of ease of use, appearance, quality of content	1-5
Initial visual-spatial and math ability tests	20 multiple-choice questions (likert scale)	1-5 per question (total 20-100)

Validation and questionnaire data were analyzed with a percentage of eligibility, i.e. P = (ΣX ÷ ΣX max) × 100%. Percentages of 81-100% were categorized as very feasible, 61-80% feasible, 41-60% less feasible, and below 41% not feasible. The internal consistency of the response questionnaire was checked with the Cronbach alpha coefficient. The test data were analyzed with the Shapiro-Wilk normality test of the score difference, the paired t-test, and the Wilcoxon-marked rating test as a comparator. The magnitude of the effect was expressed with Cohen's *d* for paired data (Cohen, 1988), and the increase relative to N-gain (Hake, 1998): ≥ 0.70 high, 0.30–0.69 moderate, and < 0.30 low.

RESULTS AND DISCUSSION

Expert Validation

Both experts rated the media as very feasible: the media expert scored 76 out of 80 (95%) and the subject matter expert 92 out of 100 (92%). Details per aspect are presented in Table 3.

Table 3. Validation results of media experts and subject matter experts

Validator	Aspects	Percentage
Media members	Display	100%
	Media engineering	83,3%
	Audio-visual communication	100%
	Benefits	100%
	Overall	95,0%
Material Expert	Material suitability	100%
	Accuracy of material substance	91,7%

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

Validator	Aspects	Percentage
	Quality of learning activities	90,0%
	Language and instructions	87,5%
	Contextuality and meaningfulness	100%
	Evaluation and feedback	83,3%
	Overall	92,0%

The media expert suggested that the scene load time be accelerated and the audio quality improved. The subject matter expert scored 3 on eight items, including block arrangement, counting, domain integration, difficulty level, character narrative, and feedback. In the analysis per scene, the subject matter expert scored 43 out of 48 (89.6%); a score of 3 was given to the Light Portal, Playground, Cheerful Kitchen, Golden Ball Quiz, and Twilight Crystal. His written advice was "it needs to be slowed down" and "the rotation is too sensitive".

Product Revision

Inputs from validation and trials and their follow-up are summarized in Table 4.

Table 4. Product input and revision

Source	Feedback	Revision
Media members	Old scene load time	Fixed in the August–October 2026 revision
Media members	Poor audio	Fixed in the August–October 2026 revision
Material Expert	"Needs to be slowed down"	Fixed in the August-October 2026 revision
Material Expert	"The turnover is too sensitive"	Fixed in the August-October 2026 revision
Small groups	The opening language and the language of the characters need to be improved	Fixed before large group tests

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

Source	Feedback	Revision
Small groups	The command is too long; needs to be shortened and simplified	Fixed before large group tests
Small groups	The screen is too crowded; the star display and the words "Perfect" are omitted	Fixed before large group tests
Small groups	Garden scene: the figure appears to be submerged on the ground	Fixed before large group tests
Small groups	Room scene: book on bed replaced doll	Fixed before large group tests
Small groups	Class scenes: too many seats, errors, too many trickster choices	Fixed before large group tests
Small groups	Market Scene: The Feet of the Character Are Bent	Fixed before large group tests
Small groups	Vima Voice Vimo	Fixed before large group tests
Large groups	All orders are delivered at once at the beginning of a scene (e.g. building houses, bridges, and towers), in almost all scenes	Planned: commands are delivered one at a time after the previous step is completed

Small Group Trials

The responses of 25 children as of October 29, 2026 were overwhelmingly positive, with a total score of 2,289 out of 2,500 (91.6%; average of 4.58 out of 5) as listed in Table 5.

Table 5. Children's response to a small-group trial (n = 25)

Aspects	Item	Score	Maximum	Percentage
Ease of use	1–3	329	375	87,7%
Display	4–10	786	875	89,8%
Content quality	11–20	1.174	1.250	93,9%
Total	**1–20**	**2.289**	**2.500**	**91,6%**

The total score per child ranged from 79–97, and the alpha Cronbach questionnaire was 0.78. Item 16 ("I can give an assessment of the results of my own work") scored 5 out of all

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

children. Observation notes showed two types of problems. First, language and instruction problems: the opening language and the language of the characters needed to be fixed, the commands were too long, and the voice of the Vima character was confused with the voice of Vimo. Second, the problem of the display per scene: the screen was too crowded, the character seemed to be immersed in the park scene, and the classroom scene contained too many chairs and the choice of tricksters and had errors.

Large Group Trials

The responses of 49 children were also very positive, with a total score of 4,655 out of 4,900 (95.0%; average 4.75) as listed in Table 6.

Table 6. Children's response to a large group trial (n = 49)

Aspects	Item	Score	Maximum	Percentage
Ease of use	1–3	676	735	92,0%
Display	4–10	1.655	1.715	96,5%
Content quality	11–20	2.324	2.450	94,9%
Total	**1–20**	**4.655**	**4.900**	**95,0%**

This score was very uniform: the total per child ranged only from 91-99 out of 100, 75.5% of the answers were 5, and there were no answers of 1 or 2. The alpha Cronbach questionnaire at this stage was -0.02, so its internal consistency could not be shown. The initial visual-spatial and math ability scores of 100 children increased after using the media (Table 7).

Table 7. Initial visual-spatial and mathematical ability scores in large group trials (n = 100)

Statistics	Loans	Posts	Differences
Average (SD)	68,05 (13,22)	79,26 (11,40)	11,21 (16,10)
Median	70,0	79,0	9,5
Range	40–100	45–100	−25 to 60

The difference in the score was normally distributed (Shapiro-Wilk, $p = 0.19$). The paired t-

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

test showed a significant improvement, $t(99) = 6.96$, $p < 0.001$, with a mean difference of 11.21 (95% IK 8.02–14.40) and $d = 0.70$. The Wilcoxon test gave the same conclusion ($p < 0.001$). The class n-gain was 0.35 (moderate). A total of 73 children experienced an increase in scores, 18 children experienced a decrease, and 9 children did not change. Three children had obtained a score of 100 on the pretest. Observation of the large group test found one major problem that appeared in almost all the scenes: the character delivered all the commands at once at the beginning. For example, the children were asked to build houses, bridges, and towers in a series of explanations before starting to work on the first step.

DISCUSSION

The validation results showed that AR-VisuMathia was feasible in terms of media and materials. The highest score was given on the suitability of the material with the child's development and the proximity of the context of the scene to daily life, which became the basis of the design since the needs analysis. The lowest scores were in media engineering (83.3%) and evaluation and feedback (83.3%). This is in line with the teacher's input at the needs analysis stage about loading time and audio, so that technical aspects and feedback are the priority of the revision.

Children's responses were very positive in both trials. However, very high scores need to be read carefully. In large group tests, almost all answers were valued at 4 or 5, and the questionnaire items did not correlate with each other. This pattern is prevalent when early childhood answers statements read by adults: younger children tend to choose the end of the scale (Chambers & Johnston, 2002), and children's understanding of number-based graded scales is still limited (Mellor & Moore, 2014). Therefore, the figures of 91.6% and 95.0% are more accurately read as signs that children are receiving and enjoying media, rather than as a measure of precise practicality.

Test scores increased with moderate to large effects ($d = 0.70$). This finding is in line with evidence that spatial exercises can improve children's abilities (Uttal et al., 2013) and that spatial activities can support numerical abilities (Hawes et al., 2017; Lowrie et al., 2017). However, the design of one pretest-postage group could not separate the influence of media from routine classroom learning, the effects of practice from taking the test twice, and the child's natural development. The presence of 18 children whose scores dropped also suggests that the benefits of media are not uniform for all children. Because the effect of spatial exercises tends to be greater when using concrete objects (Hawes et al., 2022), AR-VisuMathia should be positioned as a

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

complement, not a substitute, for activities with real objects.

The most consistent finding across the stage was the instruction load. Subject matter experts suggested that the presentation be slowed down, the small group test recorded too long commands, and the large group test found that the figure delivered all the commands at once. When a lot of important information is presented in a row, the learner's working memory capacity is easily exceeded; presentation in small segments that learners follow has been shown to result in deeper learning (Mayer & Pilegard, 2014). For children aged 5-8 years, oral instructions need to be delivered one step at a time, with the next step given after the previous step is completed. Revisions to the results of the small group test, including shortening commands, have already been implemented before the large group test. This problem persists in the large group test, which indicates that shortening each command is not enough as long as all the commands are still delivered at once at the beginning of the scene. What needs to be changed is the order in which they are delivered, not just the length of the sentence.

Practically, this trial showed that the shared screen model could be run in a classroom with a large number of children, as per the finding that the impact of the classroom screen depends on how the teacher guides the interaction (Higgins et al., 2007; Kennewell & Beauchamp, 2007). Children's visual comfort during the wearing of anaglyph glasses has not been systematically measured, whereas visual discomfort is a known risk in stereoscopic displays (Lambooij et al., 2009; McIntire et al., 2014).

This research has several limitations, namely:

1. Validation is carried out by one media expert and one material expert. The subject matter expert assesses on August 10, 2026, before the total revision, while the sheet is signed in October 2026 after the revision is completed.
2. The new expert validation includes the application components and their materials, while the teacher guide has not been validated separately.
3. The response questionnaire was adapted from an instrument for older students; some items, especially items 14-18 on writing and grading one's own work, were not yet fully suited to the abilities of children aged 5-8 years.
4. In the small group test, the questionnaire is accompanied by different parties (teachers and researchers), so the way the statement is read may not be uniform, and the presence of media developers can encourage more positive answers.
5. In the large group test, the questionnaire was only filled with 49.

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

6. The child's visual comfort and behavior during sessions have not been observed with structured instruments.

CONCLUSION

AR-VisuMathia was rated very feasible by media experts (95%) and subject matter experts (92%), was very well received by children in small and large group trials, and could be run by teachers in the classroom as a shared screen medium. Children's initial visual-spatial and math ability scores increased after using the media, but this improvement could not be attributed to the media because there were no comparison groups.

Follow-up research needs to test the effectiveness of the media with a quasi-experimental design involving a comparison group, using standardized and equivalent tests between pretests and postes, and measuring children's responses with behavioral observations and illustrated scales appropriate for 5-8 years of age. Children's visual comfort and self-efficacy in spatial-mathematical tasks also need to be systematically measured.

REFERENCES

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27–36.

Clements, D. H., & Sarama, J. (2011). Early childhood mathematics intervention. Science, 333(6045), 968–970. https://doi.org/10.1126/science.1204537

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. Developmental Psychology, 43(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66(1), 64–74. https://doi.org/10.1119/1.18809

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. Cognition and Instruction, 35(3), 236–264. https://doi.org/10.1080/07370008.2017.1323902

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

Hawes, Z. C. K., Gilligan-Lee, K. A., & Mix, K. S. (2022). Effects of spatial training on mathematics performance: A meta-analysis. Developmental Psychology, 58(1), 112–137. https://doi.org/10.1037/dev0001281

Higgins, S., Beauchamp, G., & Miller, D. (2007). Reviewing the literature on interactive whiteboards. Learning, Media and Technology, 32(3), 213–225. https://doi.org/10.1080/17439880701511040

Insani, Z., & Firdaus, F. M. (2024). Development of Augmented Reality (AR)-based space-building applications to improve spatial intelligence and self-regulated learning. Journal of Obsession: Journal of Early Childhood Education, 8(5), 1185–1196. https://doi.org/10.31004/obsesi.v8i5.6141

Ministry of Education, Culture, Research, and Technology. (2024). Regulation of the Minister of Education, Culture, Research, and Technology Number 12 of 2024 concerning the Curriculum in Early Childhood Education, Primary Education, and Secondary Education. Ministry of Education and Culture.

Kennewell, S., & Beauchamp, G. (2007). The features of interactive whiteboards and their influence on learning. Learning, Media and Technology, 32(3), 227–241.

Lambooij, M., IJsselsteijn, W., Fortuin, M., & Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. Journal of Imaging Science and Technology, 53(3), 030201. https://doi.org/10.2352/J.ImagingSci.Technol.2009.53.3.030201

Lowrie, T., Logan, T., & Ramful, A. (2017). Visuospatial training improves elementary students' mathematics performance. British Journal of Educational Psychology, 87(2), 170–186. https://doi.org/10.1111/bjep.12142

Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (2nd ed., pp. 316–344). Cambridge University Press.

McIntire, J. P., Havig, P. R., & Geiselman, E. E. (2014). Stereoscopic 3D displays and human performance: A comprehensive review. Displays, 35(1), 18–26.

Mellor, D., & Moore, K. A. (2014). The use of Likert scales with children. Journal of Pediatric Psychology, 39(3), 369–379. https://doi.org/10.1093/jpepsy/jst079

Nurillah, D., Utami, A. D., Pujiastuti, S. I., Ahdhianto, E., Nurhamdayani, & Carmenita, R. (2026). AR-VisuMathia: Developing a shared-screen 3D adventure model for visual-spatial and early mathematics skills in children aged 5–8 [Manuscript in the process of publication].

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License

OECD. (2026). PISA 2025 results (Volume I): Future-ready students. OECD Publishing. https://doi.org/10.1787/73451bc5-en

Peker, T. V., Akarca Dizakar, S. Ö., & Dayanır, D. (2022). Developing an anatomy lecture model for medical students using 3D anaglyph imaging method. Gazi Medical Journal, 33(4), 356–359. https://doi.org/10.12996/gmj.2022.80

Rakhmawati, Y., & Mustadi, A. (2022). The circumstances of literacy numeracy skill: Between notion and fact from elementary school students. Journal of Prima Edukasia, 10(1). https://doi.org/10.21831/jpe.v10i1.36427

Read, J. C. A., & Bohr, I. (2014). User experience while viewing stereoscopic 3D television. Ergonomics, 57(8), 1140–1153. https://doi.org/10.1080/00140139.2014.914581

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. Psychological Bulletin, 139(2), 352–402. https://doi.org/10.1037/a0028446

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. Child Development, 85(3), 1062–1076. https://doi.org/10.1111/cdev.12165

Copyright at authors some right reserved this work is licensed under a Creative Common
s Attribution-Share Alike 4.0 International License