



Validation of a 3d augmented reality instrument for cell biology learning

Asrizal Wahdan Wilsa^{1,2*}, Hartono³, Bambang Subali⁴, Enni Suwarsi Rahayu⁵

¹ Doctoral Program in Natural Science Education, FMIPA, Universitas Negeri Semarang, Indonesia

² Elementary School Teacher Education Program, Universitas Darul Ma'arif, Indonesia

³ Natural Science Education Program, FMIPA, Universitas Negeri Semarang, Indonesia

⁴ Physics Education Program, FMIPA, Universitas Negeri Semarang, Indonesia

⁵ Biology Program, FMIPA, Universitas Negeri Semarang, Indonesia

*Corresponding author: asrizal@rocketmail.com

ARTICLE INFO

Article history

Received: 16 August 2026

Revised: 19 September 2026

Accepted: 06 October 2026

Keywords:

Augmented reality

Cell biology

Content validity

Fuzzy delphi

Visuospatial ability

ABSTRACT

Cell biology involves abstract structures and processes that are difficult to observe directly and requires students to construct accurate microscopic and visuospatial representations. Augmented Reality can provide three-dimensional visualization of cellular structures; however, assessment instruments for AR-supported learning require careful validation. This study examined the content validity of an instrument developed for a 3D AR-integrated cell biology e-module and refined its indicators through expert discussion. The instrument comprised four components: learning outcomes, instructional content, learning processes, and learning assessment, with indicators related to microscopic representation and visuospatial ability. The study employed a research and development design guided by the ADDIE model. The study focused on the Analysis, Design, and Development stages, while the Implementation and Evaluation stages were not conducted; therefore, empirical testing with students was beyond the scope of this study. Five experts in cell biology, educational technology, and elementary education participated in the validation. Quantitative validation used the Fuzzy Delphi method with triangular fuzzy numbers and a consensus criterion of $d \leq 0.20$, while qualitative refinement was conducted through Focus Group Discussion. All items fulfilled the consensus criterion, with threshold values ranging from 0.140 to 0.199. The FGD identified areas for refinement, particularly indicator wording, representational translation, visuospatial assessment, and cellular processes. The findings provide evidence of content validity based on expert consensus and pedagogical refinement. The resulting instrument can serve as a preliminary assessment tool for microscopic representation and visuospatial ability in AR-supported cell biology learning, although further empirical testing is required to establish its psychometric properties.

© 2026 Universitas Negeri Jakarta. This is an open-access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0>)

INTRODUCTION

Cell biology is one of the basic subjects in the biology curriculum. However, learning cell biology presents challenges because many cell structures and processes are microscopic and cannot be observed directly (Gonçalves & Oliveira, 2025; Uminski et al., 2025; Wibowo et al., 2026). In the context of elementary education, cell biology should be understood as foundational biological knowledge rather than as a standalone subject entitled “cell biology.” The Indonesian elementary school curriculum integrates biological concepts within the broader Science and Social Studies (IPAS) learning framework (Badan Standar, Kurikulum, dan Asesmen Pendidikan [BSKAP], 2024). Therefore, the cell biology content addressed in this study serves as foundational knowledge for pre-service elementary school teachers who need to understand biological structures and processes to explain science concepts appropriately at the elementary level.

Thus, cell biology education requires students to develop mental models that integrate structural properties, spatial relations, and biological functions (Bouali et al., 2026). When these representations are insufficient, students may rely on verbal descriptions or static illustrations and fail to build an integrated understanding of cellular organization and processes (Hasanah et al., 2025). The ability to read and translate biological information between visual, spatial, and symbolic representations is therefore crucial to meaningful learning in cell biology (Ainsworth, 2006; Treagust & Tsui, 2013).

In this context, the importance of microscopic representation and visuospatial ability is the students' need to interpret the relative position, structure, orientation, and function of cellular components. Students can struggle to link cellular morphology to biological processes due to challenges in visualizing three-dimensional structures (Owusu et al., 2025; Gonçalves & Oliveira, 2025). Hence, learning materials for cell biology should provide learners with the opportunity to observe, manipulate, interpret, and translate representations, rather than simply providing information in verbal or two-dimensional forms.

The development of digital technologies gave rise to new ways of solving problems. One of them is Augmented Reality (AR). AR is an interactive technology that enables learners to interact with 3D representations by superimposing virtual objects in their real-world environment (Madrid & Herrero, 2025). Such technology has specific uses in the teaching of biology, as students can see the structure of cells from different perspectives and manipulate it in ways that are hard to reproduce through conventional print-based resources (Vardar-Ulu et al., 2024). AR has been shown to improve conceptual understanding, visualization, learner engagement, and spatial learning (Akçayır & Akçayır, 2017; Ibáñez & Delgado-Kloos, 2018). In biology learning, augmented reality (AR) has been used to support learners' understanding of biological systems and cell structures (Natasha et al., 2026; Rini et al., 2022; Triyanto et al., 2025). These developments are evidence of AR's growing strength as a powerful learning technology that helps students visualize and understand complex biological phenomena.

However, integrating AR into biology education does not guarantee meaningful learning. Three-dimensional visualization yields richer representations. However, students still need to interpret those representations, link them to biological concepts, and translate visual information into adequate explanations. Thus, the effectiveness of an AR-supported learning environment is determined not only by the quality of the learning media but also by the quality of the instruments used to measure the competencies that the learning environment is intended to develop (Arshad et al., 2026; Henne et al., 2024). When the representation of instruction and assessment is mismatched, the students' actual learning may be inadequately assessed.

This issue remains insufficiently addressed in research on AR-based biology learning. Much of the existing literature has focused on the development of AR media, students' learning results, perceptions, and usability. For example, some recent publications have reviewed AR learning media and its role in biology learning, and systematic reviews have reported the increasing use of AR in biological education (Indrati & Masing, 2025; Natasha et al., 2026). Nevertheless, relatively little attention has been paid to the development and validation of assessment instruments designed to capture students' microscopic representations and visuospatial abilities in an AR-supported cell biology learning environment. This creates an important methodological gap because the instructional affordances of three-dimensional visualization should be accompanied by assessment indicators that capture the representational competencies it is expected to support.

Instrument development for technology-enhanced learning also requires careful attention to

content validity. Expert judgment is often used to assess the relevance, clarity, and consistency of the assessment indicators with the intended constructs (Djiwandono & Ginting, 2025; Setiawan et al., 2024). However, the ambiguity of linguistic responses and differences in experts' interpretations of the importance of individual indicators may influence conventional expert judgment. An alternative is the Fuzzy Delphi method, which represents linguistic judgments as triangular fuzzy numbers and combines experts' opinions while accounting for the uncertainty in those judgments (Hsu & Sandford, 2007; Ishikawa et al., 1993). The method has been applied in educational research to establish expert consensus on complex, multidisciplinary constructs.

Nevertheless, quantitative consensus alone may not sufficiently reveal pedagogical and contextual issues in an assessment instrument. An item may reach an acceptable degree of numerical consensus but nonetheless need refinement in wording, instructional context, or how a construct is reflected in learning activities (Jordan et al., 2023). Thus, a Focus Group Discussion (FGD) can provide complementary qualitative evidence by enabling experts to discuss the meaning, clarity, feasibility, and pedagogical relevance of the proposed indicators (Krueger & Casey, 2014). The combination of quantifying expert consensus and qualitative expert reflection can provide a more comprehensive basis for refining instruments for technology-enhanced learning (Grant & Khodyakov, 2025; Oxley et al., 2025).

The present study intends to fill this gap by validating an assessment instrument developed for a 3D AR-integrated cell biology e-module. The instrument was developed to assess learning outcomes associated with microscopic representation and visuospatial ability. The validation process combined the Fuzzy Delphi method to establish quantitative expert consensus with FGD to identify pedagogical and representational aspects requiring refinement. The study was guided by the following research question: To what extent does the proposed instrument demonstrate content validity and pedagogical relevance for assessing microscopic representation and visuospatial ability in 3D AR-supported cell biology learning?

The study contributes to biology education in two ways. First, it provides empirical evidence of expert-based content validity for an instrument specifically designed for three-dimensional, representation-oriented cell biology learning. Second, it demonstrates how quantitative expert consensus can be complemented by qualitative discussion to refine assessment indicators before empirical implementation. Rather than treating AR merely as a technological feature of an e-module, this study positions three-dimensional visualization as part of a representation-based learning environment in which assessment should capture students' ability to interpret and construct biological representations.

METHODS

Research Design

This study employed a research and development (R&D) design using the ADDIE framework (Branch, 2010) to develop and validate an assessment instrument for a 3D Augmented Reality (AR)-integrated cell biology e-module. The ADDIE framework consists of five stages: Analysis, Design, Development, Implementation, and Evaluation. In the present study, the research activities focused on the Analysis, Design, and Development stages because the primary objective was to establish expert-based content validity of the assessment instrument before empirical implementation.

In the Analysis stage, the learning requirements in cell biology were identified, particularly the need to represent microscopic structures, spatial relationships, and biological processes through three-dimensional visualization. The analysis also considered the learning objectives of the 3D AR-integrated cell biology e-module and the assessment requirements for microscopic representation and visuospatial ability. In the design stage, translated the identified learning requirements into the assessment instrument's structure. This stage involved determining the assessment components, developing indicators, aligning them with the e-module's learning objectives, and preparing item specifications for expert validation. The resulting instrument covered learning outcomes, instructional content, learning processes, and learning assessment. In the Development stage, the assessment items were developed and reviewed for content relevance and alignment with the intended learning outcomes. The developed instrument was then subjected to expert validation using the Fuzzy Delphi method. Qualitative refinement was conducted through Focus Group Discussion (FGD) to obtain expert feedback concerning the clarity, pedagogical relevance, representational aspects, and wording of the indicators.

The Implementation and Evaluation stages were not conducted in the present study. The instrument was not administered to students for empirical testing. Consequently, the present study provides evidence of expert-based content validity but does not provide empirical evidence of construct validity, internal consistency, reliability, or measurement sensitivity. These aspects require further investigation through subsequent field implementation and evaluation with the intended users.

Participants

Five experts participated in the instrument validation. The validators were purposively selected for their expertise and experience in the development of 3D AR-supported cell biology learning. Rather than assigning one validator to one specific validation area, all five experts independently evaluated the complete set of instrument items across all components. Thus, each item was assessed by five experts, with their diverse areas of expertise providing complementary perspectives on biological content, pedagogy, educational technology, elementary education, and assessment. Their areas of expertise are presented in Table 1. The multidisciplinary composition of the panel was intended to provide complementary judgments on biological content, pedagogical appropriateness, technology integration, and assessment.

Table 1

Validator Characteristics (role in validation: Evaluated all instrument components and items)

Validator	Area of Expertise	Relevant Experience
V1	Cell Biology	Cell biology teaching and research
V2	Biology Education	Biology education and learning assessment
V3	Educational Technology	Digital learning and instructional technology
V4	Elementary Education	Elementary science education and pedagogy
V5	Biology Education/Assessment	Biology learning and educational assessment

Note. All five experts independently evaluated all instrument components and items. The area of expertise indicates each validator's primary expertise and does not restrict them to a particular validation component.

Instrument

The instrument consisted of 38 items developed from the learning objectives of the 3D AR-integrated cell biology e-module, with emphasis on microscopic representation and visuospatial ability. The indicators covered learning outcomes (n=10), instructional content (n=14), learning processes (n=9), and assessment (n=5). Experts rated each item using a seven-point linguistic agreement scale. The linguistic responses were transformed into triangular fuzzy numbers (l, m, u). Consensus was considered achieved when the threshold value d was ≤ 0.20 and the proportion of experts reaching consensus met the predetermined criterion. Defuzzification values were used to describe the strength of the aggregated expert judgment. The complete item structure is presented in Table 2.

Table 2

Instrument Components and Item Distribution

Component	Indicator Focus	Item Codes	Number of Items
Learning Outcomes	Microscopic representation and visuospatial learning outcomes.	B.1–B.10	10
Instructional Content	Cellular structures, spatial relationships, functions, and biological representations.	C.1.1–C.4.3	14
Learning Process	Observation, manipulation, interpretation, and representational translation.	D.1–D.3.3	9
Learning Assessment	Representation-oriented and visuospatial assessment.	E.1.1–E.2.3	5
Total			38

Data Collection and Analysis

Experts assessed each item using a linguistic agreement scale. The responses were converted into triangular fuzzy numbers (TFNs) and analyzed using the Fuzzy Delphi method. Expert consensus was determined using a threshold value of $d \leq 0.20$, followed by defuzzification to obtain the aggregated judgment for each item. The 3D AR-integrated cell biology e-module and the assessment instrument developed in this study are available at <https://heyzine.com/flip-book/49ffa4a700.html>.

The validated items were subsequently discussed through an FGD to identify aspects requiring pedagogical and representational refinement. The data from the FGD were thematically analyzed, focusing on the clarity of indicators, representational activities, visuospatial assessment, and AR integration. The final instrument was revised by integrating the quantitative Fuzzy Delphi results and qualitative FGD findings.

In addition, document analysis was conducted to examine the alignment between the learning objectives of the 3D AR-integrated cell biology e-module and the assessment instrument's components, indicators, and items. The analysis focused on whether each assessment component and indicator aligned with the intended learning outcomes and learning activities in the e-module. The findings from the document analysis were compared with the Fuzzy Delphi results and FGD feedback to identify potential inconsistencies and support the refinement of the instrument. Thus, methodological triangulation in this study involved three complementary sources of evidence: quantitative expert consensus obtained through the Fuzzy Delphi method, qualitative expert feedback obtained through the FGD, and document analysis of the alignment between the e-module and the assessment instrument.

RESULTS AND DISCUSSION

Fuzzy Delphi Validation Results

The Fuzzy Delphi analysis was conducted on 38 items of the 3D Augmented Reality (AR)-integrated cell biology learning instrument. All items achieved the predetermined consensus criterion of $d \leq 0.20$. The threshold values ranged from 0.140 to 0.199, indicating that all items reached the required level of expert agreement. The highest threshold value was obtained for item B.1 ($d = 0.199$), whereas several items obtained the lowest threshold value of $d = 0.140$. Although the degree of agreement varied across items, no item exceeded the predetermined threshold. Therefore, all 38 items were retained for further qualitative examination.

The decision criterion was established before the validation analysis. An item was considered to have achieved expert consensus when the threshold value (d) was ≤ 0.20 . Items with $d > 0.20$ would require further review and revision based on the experts' qualitative feedback. In the present study, all validated items met the acceptance criterion; therefore, no item was excluded based on the Fuzzy Delphi consensus criterion. The defuzzification results also indicated positive expert judgments across the validated items. The detailed values of the expert assessments, triangular fuzzy numbers, defuzzification values, threshold values, and item decisions are presented in [Table 3](#).

Table 3
Fuzzy Delphi Validation Results

Item Code	Item Category	Validators					Mean TFN			Defuzzification	Threshold	Decision
		V1	V2	V3	V4	V5	<i>l</i>	<i>m</i>	<i>u</i>	A		
B.1	Learning Outcomes	7	6	4	6	6	0.634	0.798	0.934	0.789	0.199	Accepted
B.2		6	7	5	6	7	0.700	0.866	0.966	0.844	0.179	Accepted
B.3		7	7	5	6	6	0.700	0.866	0.966	0.844	0.179	Accepted
B.4		7	7	5	6	7	0.732	0.900	0.966	0.866	0.173	Accepted
B.5		7	6	6	6	7	0.734	0.898	1.000	0.877	0.169	Accepted
B.6		6	6	6	6	6	0.670	0.830	1.000	0.833	0.163	Accepted
B.7		7	7	6	6	7	0.766	0.932	1.000	0.899	0.159	Accepted
B.8		7	7	7	6	7	0.798	0.966	1.000	0.921	0.140	Accepted
B.9		6	7	7	6	6	0.734	0.898	1.000	0.877	0.169	Accepted
B.10		7	7	7	7	6	0.798	0.966	1.000	0.921	0.140	Accepted
C.1.1	Instructional Content	6	6	5	7	6	0.668	0.832	0.966	0.822	0.176	Accepted
C.1.2		7	7	7	7	6	0.798	0.966	1.000	0.921	0.140	Accepted
C.1.3		6	7	7	7	5	0.732	0.900	0.966	0.866	0.173	Accepted
C.2.1		6	6	7	7	6	0.734	0.898	1.000	0.877	0.169	Accepted
C.2.2		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
C.2.3		7	7	5	7	6	0.732	0.900	0.966	0.866	0.173	Accepted

Item Code	Item Category	Validators					Mean TFN			Defuzzification	Threshold	Decision
		V1	V2	V3	V4	V5	<i>l</i>	<i>m</i>	<i>u</i>	A		
C.2.4		7	7	5	7	6	0.732	0.900	0.966	0.866	0.173	Accepted
C.2.5		7	7	6	7	7	0.798	0.966	1.000	0.921	0.140	Accepted
C.3.1		7	7	5	7	6	0.732	0.900	0.966	0.866	0.173	Accepted
C.3.2		7	7	5	7	6	0.732	0.900	0.966	0.866	0.173	Accepted
C.3.3		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
C.4.1		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
C.4.2		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
C.4.3		7	7	7	7	6	0.798	0.966	1.000	0.921	0.140	Accepted
D.1		7	7	5	7	6	0.732	0.900	0.966	0.866	0.173	Accepted
D.2.1	Learning Process	7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.2.2		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.2.3		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.2.4		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.2.5		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.3.1		7	7	7	6	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.3.2		7	7	7	6	6	0.766	0.932	1.000	0.899	0.159	Accepted
D.3.3		7	7	7	6	6	0.766	0.932	1.000	0.899	0.159	Accepted
E.1.1	Learning Assessment	7	7	6	6	6	0.734	0.898	1.000	0.877	0.169	Accepted
E.1.2		7	7	6	6	6	0.734	0.898	1.000	0.877	0.169	Accepted
E.2.1		7	7	7	6	6	0.766	0.932	1.000	0.899	0.159	Accepted
E.2.2		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted
E.2.3		7	7	6	7	6	0.766	0.932	1.000	0.899	0.159	Accepted

Focus Group Discussion Results

The FGD was conducted after the quantitative validation to obtain expert feedback on the instrument's clarity, pedagogical relevance, and representational characteristics. Four major areas of refinement emerged from the discussion. First, several indicator statements were simplified and reformulated to make them more operational and reduce ambiguity. Second, the experts recommended strengthening representational translation activities by requiring students to connect three-dimensional visualizations with conceptual explanations. Third, the assessment indicators were refined to place greater emphasis on visual representation and visuospatial abilities rather than factual identification alone. Fourth, the experts recommended extending AR activities beyond visualizing cellular structures to include selected cellular processes.

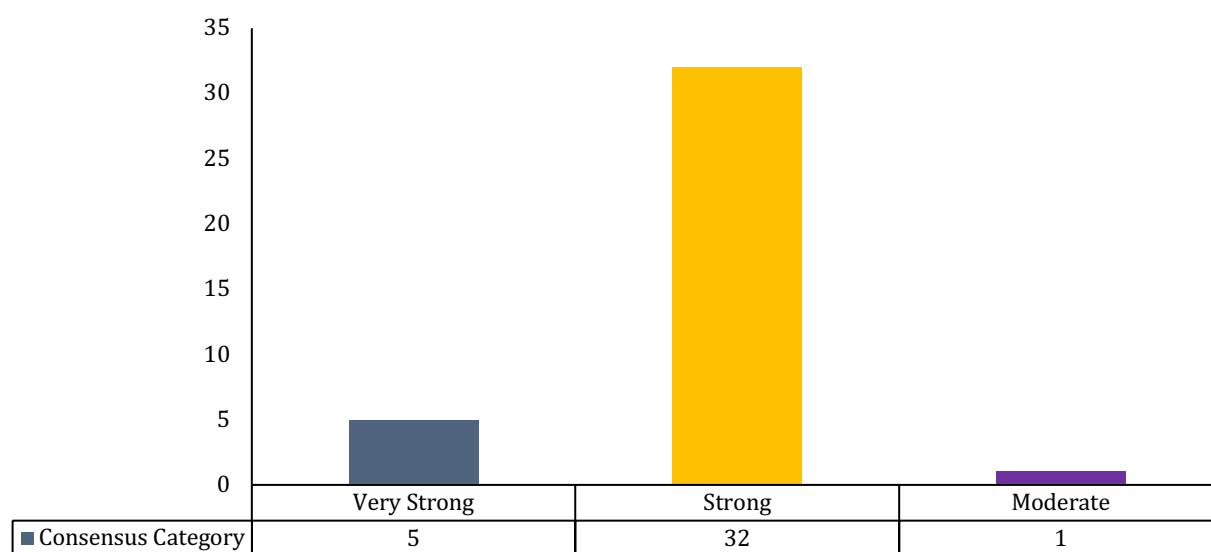


Figure 1. 3D AR Integration in the Cell Biology E-Module

The FGD did not result in the rejection of any item that had met the Fuzzy Delphi consensus criterion. Instead, the discussion provided recommendations for improving the wording, pedagogical relevance, and implementation of the accepted indicators.

Content Validity of the Instrument

The Fuzzy Delphi findings provide evidence of strong expert agreement regarding the content relevance of the 38 indicators developed for the 3D AR-integrated cell biology learning instrument. All items achieved the predetermined threshold criterion ($d \leq 0.20$), with values ranging from 0.140 to 0.199. The relatively narrow range indicates that the validators generally shared similar judgments regarding the relevance and clarity of the proposed indicators. This finding is important because the instrument was designed to assess competencies that extend beyond factual knowledge, particularly microscopic representation and visuospatial ability.

The highest threshold value was observed for item B.1 ($d = 0.199$). Although this item still satisfied the consensus criterion, its relatively higher threshold indicates greater variation in expert judgments compared with other indicators. Rather than interpreting this finding as evidence of item weakness, it provides an important basis for qualitative examination. The subsequent FGD therefore served to clarify the meaning and pedagogical application of the indicators. This illustrates the advantage of combining quantitative consensus with qualitative expert reflection in instrument development.

The content validity demonstrated in this study is particularly relevant to cell biology learning. The cellular frameworks and processes are conceptually complex since learners have to understand entities that are too small to be directly observed and deduce their organization thru representations (Windyariani et al., 2023). Therefore, the cell biology assessment tools should measure students' ability to analyze the connections between structure, spatial orientation, and biological function, aside from conventional factual questions (Walton, 2023; Wollmuth et al., 2023).

The finding also has relevance to the current development of AR in biology education. Recent research by Natasha et al. (2026), Rini et al. (2022), and Triyanto et al. (2025) has demonstrated the increasing application of AR for visualizing biological concepts, including cellular structures and human biological systems. A recent systematic review also pointed out that the research in AR in biology education has increased substantially with cell-related topics being among the frequently investigated domains. However, the measurement of learning outcomes in AR research has often been carried out thru traditional tests, which means that there is still a need to develop assessment tools that are specifically tailored to three-dimensional representation and visuospatial skills (Indrati & Masing, 2025; Setiawan et al., 2025).

Therefore, the contribution of the present study is not simply the production of another AR-related instrument. Its contribution lies in aligning assessment with the specific representational affordances of 3D AR. The instrument was designed to assess competencies that become particularly meaningful when learners interact with three-dimensional biological objects, including interpreting spatial relationships, transforming visual information, and connecting representations with biological explanations.

Pedagogical and Representational Refinement through FGD

Although the Fuzzy Delphi results indicated that all items achieved consensus, the FGD revealed that numerical agreement alone was insufficient to ensure that every indicator was pedagogically optimal. Four major areas were identified for refinement: indicator wording, representational translation, visuospatial assessment, and the integration of cellular processes. The significance of MSA refinement of the indicator wording is that assessment items should effectively convey the intended cognitive action. The general construct may be relevant for the experts, but students may interpret the ambiguous indicators differently (Susanto et al., 2023; Zeng et al., 2024). Therefore, some indicators have been rephrased in more operational statements by the FGD. This refinement improves the link between the intended competency and the observable learning performance (Hjellvik et al., 2024).

More importantly, the FGD contributed by its focus on representational translation. In the present study, the experts recommended that students should not only identify or observe three-dimensional cell structures but also be asked to connect visual representations with conceptual explanations. This recommendation reflects the pedagogical judgment of the expert validators and does not constitute empirical evidence of students' actual performance. Such an approach is consistent

with previous research emphasizing the importance of connecting visual representations with conceptual explanations (Edelsbrunner & Hofer, 2024; Geçişler et al., 2025; Lore et al., 2024). It is also consistent with the multiple representations perspective, which suggests that meaningful learning involves coordinating information from multiple representations rather than processing each representation independently (Ainsworth, 2006).

In the context of the present instrument, representational translation is particularly relevant because cell biology involves multiple forms of representation, including diagrams, microscopic images, three-dimensional models, symbolic descriptions, and verbal explanations. The ability to translate between such representations has been discussed in previous research as an important aspect of meaningful biological understanding (Thyberg et al., 2025). In the present study, the FGD findings support the inclusion of representational translation as an assessment component; however, whether students can effectively perform such translation cannot be established from the present expert validation and requires subsequent empirical testing.

The FGD also shifted the emphasis of the instrument from factual identification toward representation-oriented performance. This is an important pedagogical distinction. A student may correctly identify a mitochondrion, nucleus, or cell membrane without being able to explain its spatial relationship with other cellular components or relate its structure to its function. An assessment instrument designed for a 3D AR environment should therefore examine not only what students know but also how they interpret and communicate what they see.

The Role of 3D AR in Supporting Biological Representation

The findings indicate that the pedagogical value of AR should not be reduced to its visual attractiveness or technological novelty. The important feature of 3D AR for cell biology is its capacity to provide interactive spatial representations that can be manipulated and examined from different perspectives.

This feature is important especially for microscopic representation, since students can see cell structures, which are not directly observable. In the same 3D environment, learners are able to manipulate static 2D images (e.g., shape, position, relative size, and spatial relationships) (Abdo et al., 2024; Wollmuth et al., 2023). AR has been successfully used in science and biology education as a means to improve the visualization and conceptual understanding (Akçayır & Akçayır, 2017; Ibáñez & Kloos, 2018). More recently, the work of Biosfer points to the use of AR as a platform to think about abstract biological ideas (Indriani et al., 2024).

But visualization is not comprehension of a concept. Students need to interpret the representation in terms of their prior knowledge. This explains the theoretical and pedagogical significance of the FGD proposal for enhancing representational translation. The AR object is educationally meaningful when students use it as a basis for reasoning about biological structures and processes (Hajama et al., 2025; Nampira et al., 2025; Thyberg et al., 2025).

Also significant is the experts' proposal to expand AR activities from structural visualization to cellular processes. Cell biology is more than simply identifying organelles. It deals with the understanding of dynamic relations like transport, energy transformation, cellular division and interaction of cellular components. Thus, the assessment tool that captures both the structural and process-oriented representations will align better with the conceptual characteristics of cell biology (Hockenberry et al., 2024; Laubscher et al., 2024; Gonçalves & Oliveira, 2025).

Implications for Assessment in Technology-Enhanced Biology Learning

The findings have broader implications for the development of assessment instruments in technology-enhanced biology learning. The increasing availability of AR and other immersive technologies creates learning environments in which students interact with representations that differ substantially from conventional textbooks. Assessment should therefore evolve alongside the instructional technology.

The present study suggests that three dimensions should be considered when developing assessments for AR-supported biology learning: construct alignment, representational alignment, and pedagogical alignment. Construct alignment ensures that the indicators measure the intended biological competencies. Representational alignment ensures that the indicators correspond to the visual and spatial affordances provided by 3D AR. The assessment should also be pedagogically aligned to reflect meaningful learning activities rather than merely testing the students' capacity to

operate the technology (Bogomolova et al., 2020; Haaranen et al., 2024; Ngema, 2024).

This is consistent with recent research relating to AR, which suggests that the development of AR is more guided by pedagogical frameworks such as TPACK, rather than treating technology as a separate component of learning design (Rutten & Truijen, 2025). It also addresses the broader trend identified in the AR systematic review, where AR research has focused mainly on learning outcomes, motivation and critical thinking. In this paper, we propose an assessment-oriented view, more precisely, the operation of competencies in a 3-D biological representation before testing for effectiveness (Faria & Miranda, 2024; Mohammadhossein et al., 2024).

Methodologically, the integration of Fuzzy Delphi, FGD, and document analysis highlights the complementarity of different kinds of evidence for instrument validation in technology-enhanced learning. Fuzzy Delphi provides structured evidence of expert consensus on the relevance of the assessment items, while the FGD offers qualitative insights into their clarity, pedagogical relevance, representational characteristics, and contextual considerations. Document analysis further supports the alignment between the learning objectives of the 3D AR-integrated cell biology e-module and the components, indicators, and assessment items. Thus, these three approaches provide complementary evidence for the refinement of the instrument (Afrizon et al., 2025; Yacob et al., 2024).

However, the results should be considered in light of the limitations of the study. Validation evidence is based on expert judgment and does not provide empirical evidence of construct validity, internal consistency or sensitivity of measurement. Hence, the instrument needs to be field-tested with students before making strong claims about students' microscopic representation or visuospatial ability. Future research should investigate these psychometric properties and study whether the instrument could differentiate differences in students' representational performance after 3D AR-supported instruction.

CONCLUSION

This study provides evidence of content validity for a 38-item instrument designed to assess microscopic representation and visuospatial ability in 3D Augmented Reality-supported cell biology learning. The fuzzy Delphi analysis revealed that all items fell within the pre-determined consensus criterion ($d \leq 0.20$) with threshold values between 0.140 and 0.199. Results showed that the experts' panel found the indicators to be relevant and sufficiently clear. The Focus Group Discussion complemented the quantitative findings by refining indicator wording, representational translation activities, visuospatial assessment, and the integration of cellular processes into the AR-supported learning environment. The combined use of Fuzzy Delphi and FGD therefore provided complementary evidence for content validation and pedagogical refinement. The resulting instrument can serve as a preliminary assessment tool for 3D AR-supported cell biology learning, particularly for activities involving microscopic representations and visuospatial processing. However, the present findings are limited to expert-based content validation. Further empirical testing with students is required to establish construct validity, reliability, and the instrument's sensitivity to differences in learning performance.

ACKNOWLEDGMENT

We would like to thank the DPA FMIPA Universitas Negeri Semarang (UNNES) 2026 for their financial support based on Contract Number 87.7.5/UN37/PPK.04/2026. This support contributed to the conduct of the research and to the preparation of this manuscript. The authors also express sincere thanks to all persons and institutions that assisted in the successful completion of this study.

REFERENCES

- Abdo, S. N., Hsu, J. L., Kapetanakis, C., Newman, D. L., Wright, L. K., & Bailey, J. (2024). An Exploration of Spatial Visualization Skills: Investigating Students' Use of 3D Models in Science Problems during Think-Aloud Interviews. *Journal of Chemical Education*, 101(9), 3624–3634. <https://doi.org/10.1021/ACS.JCHEMED.3C01355>
- Adara Natasha, K., Anwar, Y., & Madang, K. (2026). TPACK-based augmented reality learning media for human excretory biology education. *Biosfer: Jurnal Pendidikan Biologi*, 19(1), 71–77. <https://doi.org/10.21009/BIOSFERJPB.58133>
- Afrizon, R., Ellany Mohtar, L., Syahrman, M., & Azmi, M. (2025). Development of the Fuzzy Delphi Instrument for Identifying Elements of the PhyARECi-STEM Learning Module. *EDUCATUM*

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198.
<https://doi.org/10.1016/J.LEARNINSTRUC.2006.03.001>
- 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>
- Arshad, B., Ishak, N. A., & Zaharudin, R. (2026). Validity and Reliability of The e-ReSP-AR Module for Inquiry-Based Biology Learning on Cellular Respiration. *Jurnal Pendidikan Sains Dan Matematik Malaysia*, 16(1), 70–84. <https://doi.org/10.37134/JPSMM.VOL16.1.6.2026>
- Badan Standar, Kurikulum, dan Asesmen Pendidikan. (2024). Capaian Pembelajaran Pada Kurikulum Merdeka. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
<https://lmsspada.kemdiktisaintek.go.id/mod/resource/view.php?id=136719&forceview=1>
- Bogomolova, K., van der Ham, I. J. M., Dankbaar, M. E. W., van den Broek, W. W., Hovius, S. E. R., van der Hage, J. A., & Hierck, B. P. (2020). The Effect of Stereoscopic Augmented Reality Visualization on Learning Anatomy and the Modifying Effect of Visual-Spatial Abilities: A Double-Center Randomized Controlled Trial. *Anatomical Sciences Education*, 13(5), 558–567.
<https://doi.org/10.1002/ASE.1941>
- Bouali, R., Agorram, B., Maskour, L., Zaki, M., Ksiksou, J., & Lidrissi-hassani, S. (2026). University students' misconceptions of cellular structures and functions. *Journal of Biological Education*, 60(1), 102–117. <https://doi.org/10.1080/00219266.2025.2452203>
- Branch, R. M. (2010). Instructional design: The ADDIE approach. *Instructional Design: The ADDIE Approach*, 1–203. <https://doi.org/10.1007/978-0-387-09506-6>
- Djiwandono, P. I., & Ginting, D. (2025). Evaluating research reports on the qualities of tests of English language skills in Indonesian schools: A systematic review. *Language Education & Assessment*, 8(1), 2237–2237. <https://doi.org/10.29140/LEA.V8N1.2237>
- Edelsbrunner, P. A., & Hofer, S. I. (2024). Examining and comparing the relation between representational competence and conceptual knowledge across four samples. *Frontiers in Education*, 9, 1459603. <https://doi.org/10.3389/FEDUC.2024.1459603>
- Faria, A., & Miranda, G. L. (2024). Augmented Reality in Natural Sciences and Biology Teaching: Systematic Literature Review and Meta-Analysis. *Emerging Science Journal*, 8(4), 1666–1685.
<https://doi.org/10.28991/ESJ-2024-08-04-025>
- Geçişler, G. A., İncelenmesi, A., Gökçe, H., Saliha, I., & Eyüboğlu, K. (2025). Fotosentez Öğretiminde Çoklu Gösterimlerin Öğrenmeye Etkilerinin Gösterimler Arası Geçişler Aracılığıyla İncelenmesi. *Buca Faculty of Education Journal*, (64), 1342–1366.
<https://doi.org/10.53444/DEUBEFD.1458350>
- Gonçalves, P. J., & Oliveira, C. A. (2025). Development, application, and evaluation of tools based on virtual and augmented realities for inclusive teaching and learning of cell membranes. *Anatomical Sciences Education*, 18(5), 462–484. <https://doi.org/10.1002/ASE.70024>
- Grant, S., & Khodyakov, D. (2025). Proposal for a critical appraisal tool for studies using the Delphi method (DCAT). *BMJ*, 391. <https://doi.org/10.1136/BMJ-2025-084509>
- Haaranen, L., Ahrenberg, L., & Hellas, A. (2024). Decades of Striving for Pedagogical and Technological Alignment. *ACM International Conference Proceeding Series*.
<https://doi.org/10.1145/3631802.3631809>
- Hajama, S., Chu, H.-E., & Treagust, D. (2025). *Monitoring student meaning-making in human anatomy and physiology: the need for multiple representations and representation construction* (p. 35). Australian Science Education Research Association.
<https://researchers.mq.edu.au/en/publications/monitoring-student-meaning-making-in-human-anatomy-and-physiology/>
- Hasanah, M. R., Zannah, F., Aufa, M. N., Nurkhaliza, G. N., Imelda, & Fenny. (2025). Analysis of Learning Needs for Cell Biology: a Study From the Perspectives of Lecturers and Students in Palangka Raya City. *Bioilmi: Jurnal Pendidikan*, 11(1), 26–32.
<https://doi.org/10.19109/BIOILMI.V11I1.28168>
- Henne, A., Syskowski, S., Krug, M., Möhrke, P., Thoms, L. J., & Huwer, J. (2024). How to Evaluate Augmented Reality Embedded in Lesson Planning in Teacher Education. *Education Sciences*,

- 14(3), 264. <https://doi.org/10.3390/EDUCSCI14030264>
- Hjellvik, S., Mallam, S., Giskeødegård, M. F., & Nazir, S. (2024). Review on Competency Assessment Instrumentation in Computer-based Simulation. *Technology, Knowledge and Learning*, 29(4), 2171–2200. <https://doi.org/10.1007/S10758-024-09735-4>
- Hockenberry, M. A., Daugird, T. A., & Legant, W. R. (2024). Cell dynamics revealed by microscopy advances. *Current Opinion in Cell Biology*, 90, 102418. <https://doi.org/10.1016/J.CEB.2024.102418>
- Hsu, C.-C., & Sandford, B. A. (2007). The Delphi Technique: Making Sense of Consensus. *Practical Assessment, Research, and Evaluation*, 12(1). <https://doi.org/10.7275/PDZ9-TH90>
- Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, 123, 109–123. <https://doi.org/10.1016/J.COMPEDU.2018.05.002>
- Indrati, D. A., & Masing, F. A. (2025). Research trends of augmented reality in biology learning: A systematic literature review from 2020-2024. *Biosfer: Jurnal Pendidikan Biologi*, 18(2), 145–159. <https://doi.org/10.21009/BIOSFERJPB.53911>
- Indriani, R. P., Kurniati, T. H., & Ristanto, R. H. (2024). Improving student's biological literacy skills using ARVi learning media. *Biosfer: Jurnal Pendidikan Biologi*, 17(1), 286–296. <https://doi.org/10.21009/BIOSFERJPB.43976>
- Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R., & Mieno, H. (1993). The max-min Delphi method and fuzzy Delphi method via fuzzy integration. *Fuzzy Sets and Systems*, 55(3), 241–253. [https://doi.org/10.1016/0165-0114\(93\)90251-C](https://doi.org/10.1016/0165-0114(93)90251-C)
- Jordan, J., Haas, M. R. C., Hickam, G., Murray, C. T., Hill, J., Cico, S. J., Wolff, M., Manthey, D. E., Wagner, J. C., & Santen, S. A. (2023). Development of a lecture evaluation tool rooted in cognitive load theory: A modified Delphi study. *AEM Education and Training*, 7(1), e10839. <https://doi.org/10.1002/AET2.10839>
- Krueger, R. A., & Casey, M. A. (2014). Focus groups: A practical guide for applied research. Fifth Edition. *Focus Groups: A Practical Guide for Applied Research*, 197.
- Laubscher, H., Loos, B., & Theart, R. P. (2024). Engage and learn: Improved learning of cellular structures using a virtual reality-based learning experience. *Computers & Education: X Reality*, 5, 100089. <https://doi.org/10.1016/J.CEXR.2024.100089>
- Lore, C., Lee, H. S., Pallant, A., & Chao, J. (2024). Using multiple, dynamically linked representations to develop representational competency and conceptual understanding of the earthquake cycle. *Computers & Education*, 222, 105149. <https://doi.org/10.1016/J.COMPEDU.2024.105149>
- Madrid, G. N., & Herrero, F. J. (2025). The impact of augmented reality on learning cell biology: a quasi-experimental study in secondary education. *Technology, Pedagogy and Education*. <https://doi.org/10.1080/1475939X.2025.2570835>
- Mohammadhossein, N., Richter, A., & Lukosch, S. (2024). Augmented Reality in Learning Settings: A Systematic Analysis of its Benefits and Avenues for Future Studies. *Communications of the Association for Information Systems*, 54(1), 29–49. <https://doi.org/10.17705/1CAIS.05402>
- Nampira, A. A., Sharma, S., & Wangchuk, D. (2025). Augmented Reality in STEM Education: Enhancing Conceptual Understanding and Learner Motivation. *Journal Emerging Technologies in Education*, 3(3), 117–127. <https://doi.org/10.70177/JETE.V3I3.3410>
- Ngema, D. (2024). A Critical Review of the Purposes and Authenticity of E-Assessment Policies and Practices. *Open Books and Proceedings*, 134–153. <https://doi.org/10.38140/OBP2-2024-10>
- Owusu, B. D., Stargell, L. A., Otto, J., & Balgopal, M. M. (2025). Undergraduate Students' Misconceptions About Protein Structure and Function Stem From Challenges in Visuospatial Reasoning. *Biochemistry and Molecular Biology Education*, 53(6), 579–587. <https://doi.org/10.1002/BMB.70008>
- Oxley, E., Nash, H. M., & Weighall, A. R. (2025). Consensus building using the Delphi method in educational research: a case study with educational professionals. *International Journal of Research & Method in Education*, 48(1), 29–43. <https://doi.org/10.1080/1743727X.2024.2317851>
- Rini, D. S., Azrai, E. P., Suryanda, A., Inayah, S. S., Khansa, A. A., & Kurnianto, M. B. (2022). Augmented reality (AR) technology on the android operating system in human respiratory system: From organ to cell. *Biosfer: Jurnal Pendidikan Biologi*, 15(1), 25–35. <https://doi.org/10.21009/BIOSFERJPB.23448>
- Rutten, N., & Brouwer-Truijen, K. (2025). Defining XR-Specific Teacher Competencies: Extending the

- DigCompEdu Framework for Immersive Education. *Trends in Higher Education 2025, Vol. 4, Page 11*, 4(1), 11. <https://doi.org/10.3390/HIGHEREDU4010011>
- Setiawan, E., Darmawan, E., Radian, M., & Alamsyah, N. (2025). A Meta-analysis: Trends in the use of augmented reality in biology learning (2016-2025). *Biosfer: Jurnal Pendidikan Biologi*, 18(2), 353–369. <https://doi.org/10.21009/BIOSFERJPB.57782>
- Setiawan, R., Wagiran, W., & Alsamiri, Y. (2024). Construction of an instrument for evaluating the teaching process in higher education: Content and construct validity. *REID (Research and Evaluation in Education)*, 10(1), 50–63. <https://doi.org/10.21831/REID.V10I1.63483>
- Susanto, H., Wedyaswari, M., & Dalimunthe, K. L. (2023). A content validity and cognitive interview to develop the HARMONI items: Instrument measuring student well-being in West Java, Indonesia. *Frontiers in Education*, 8, 1132031. <https://doi.org/10.3389/FEDUC.2023.1132031>
- Thyberg, A., Schönborn, K., & Gericke, N. (2025). Phases of Progression: Students' meaning-making of Epigenetic Visual Representations within and between Levels of Organization. *Research in Science Education*, 55(2), 399–423. <https://doi.org/10.1007/S11165-024-10196-Z>
- Treagust, D. F., & Tsui, C. (2013). The role of multiple external representations in learning biology. In *Multiple representations in biology education* (pp. 1–24). Springer. https://doi.org/10.1007/978-94-007-5157-1_1
- Triyanto, S. A., Bilbina, K., & Putri, R. S. (2025). The potential of 3d augmented reality book-based cell learning media to support educational transformation. *Biosfer: Jurnal Pendidikan Biologi*, 18(1), 61–68. <https://doi.org/10.21009/BIOSFERJPB.47597>
- Uminski, C., Newman, D. L., & Wright, L. K. (2025). Probing Visual Literacy Skills Reveals Unexpected Student Conceptions of Chromosomes. *CBE Life Sciences Education*, 24(1), ar17. <https://doi.org/10.1187/CBE.24-07-0176>
- Vardar-Ulu, D., Ragab, S. E., Agrawal, S., & Dutta, S. (2024). Using augmented reality in molecular case studies to enhance biomolecular structure-function explorations in undergraduate classrooms. *Journal of Microbiology & Biology Education*, 25(2). <https://doi.org/10.1128/JMBE.00019-24>
- Walton, K. L. W. (2023). Use of a short, in-class drawing activity to assess student understanding of core concepts of the cell membrane in an undergraduate physiology course. *Advances in Physiology Education*, 47(3), 508–513. <https://doi.org/10.1152/ADVAN.00218.2022>
- Wibowo, J., Hindriana, A. F., & Abidin, Z. (2026). Metaphor-Based Visual Media for Cell Biology Instruction: Effects on Students' Science Literacy and Metacognitive Skills. *Tadris: Jurnal Keguruan Dan Ilmu Tarbiyah*, 11(1), 146–158. <https://doi.org/10.24042/TADRIS.V11I1.28912>
- Windyarani, S., Sabilla, A. D., Ramdhan, B., & Irawan, B. (2023). Mental Representation Level of Junior High School Students in Cell Biology Concept Based on Gender. *Jurnal Pendidikan IPA Indonesia*, 12(3), 377–386. <https://doi.org/10.15294/JPII.V12I3.42316>
- Wollmuth, E. M., Correa, A., Alvarado Obando, M., Smith, M. K., Buckley, D. H., Hefferon, K. L., & Angert, E. R. (2023). Helping students see bacteria in 3D: cellular models increase student learning about cell size and diffusion. *Journal of Microbiology & Biology Education*, 24(3). <https://doi.org/10.1128/JMBE.00089-23>
- Yacob, N. S., Yunus, M. M., & Hashim, H. (2024). A Quest for Experts' Consensus on the Knowledge Domain for an ESL Global Competence Framework Using Fuzzy Delphi Method Analysis. *International Journal of Learning, Teaching and Educational Research*, 23(11), 1–31. <https://doi.org/10.26803/IJLTER.23.11.1>
- Zeng, B., Jeon, M., & Wen, H. (2024). How does item wording affect participants' responses in Likert scale? Evidence from IRT analysis. *Frontiers in Psychology*, 15, 1304870. <https://doi.org/10.3389/FPSYG.2024.1304870>