ITEJ
Indonesian Technology and Education Journal
E-ISSN: 2985-8216; P-ISSN: 2985-8208
https://journal.diginus.id/index.php/ITEJ

Learning Outcomes Following Assemblr EDU-Based Augmented Reality Instruction on the Human Respiratory System

Murti Handayani[1], Nurhayani H. Muhiddin[2], Rifda Nurhikmawati Arif[3]
[1,2,3]*Universitas Negeri Makassar, Makassar, Indonesia*

Corresponding e-mail : handayanimurtiars25@gmail.com

ARTICLE INFO

Keywords:
Assemblr EDU;
Augmented Reality;
Human Respiratory System;
Learning Outcomes;
Science Education.

Article History

Received: April 10, 2026
Revised: July 06, 2026
Accepted: August 08, 2026

ABSTRACT

The human respiratory system involves complex anatomical structures and physiological processes that are difficult to observe directly, creating challenges for students in developing conceptual understanding. This study aimed to describe changes in middle school students' learning outcomes following instruction supported by Assemblr EDU-based augmented reality on the human respiratory system. A quantitative pre-experimental approach with a one-group pretest–posttest design was employed. Participants were selected through purposive sampling, and learning outcomes were assessed using a multiple-choice achievement test covering remembering, understanding, applying, and analyzing cognitive processes. The data were analyzed descriptively by comparing pretest and posttest performance and evaluating normalized learning gains. The findings showed that students achieved substantially higher learning outcomes after participating in augmented reality-supported instruction, with overall learning gains categorized as high. Improvements were also observed across all cognitive-process categories, although differences among these categories should be interpreted cautiously because of unequal item distribution. These findings indicate that Assemblr EDU can serve as a useful three-dimensional representational resource for supporting the learning of internal anatomical structures when integrated with appropriate explanations and cognitive learning activities. However, because the study did not include a comparison group or inferential analysis, the findings should be interpreted as preliminary evidence of learning changes rather than causal evidence of instructional effectiveness.

To cite this article : Handayani, M., Muhiddin, N. H., Arif, R. N. (2026). Learning Outcomes Following Assemblr EDU-Based Augmented Reality Instruction on the Human Respiratory System. *Indonesian Technology and Education Journal, 4*(2), 365-377. Doi. 10.61255/itej.v4i2.1610

INTRODUCTION

Science learning requires students to remember terms and facts, understand relationships among concepts, apply concepts in specific situations, and analyze cause-and-effect relationships within a system. These demands are particularly important at the middle school level, when students begin to move from factual knowledge toward more complex conceptual reasoning. The selection of instructional media should therefore consider the characteristics of the content, the learning objectives, and the cognitive processes to be developed. Visually appealing media do not necessarily promote deeper understanding unless they are integrated into purposeful learning activities. Technology should be positioned as part of a pedagogical design that supports teachers' explanations and students' thinking rather than replacing either.

The human respiratory system is a topic that requires the integration of factual, conceptual, and procedural knowledge. Students must identify the respiratory organs, explain the function of each organ, understand the pathway of inhaled and exhaled air, distinguish the mechanisms of inspiration and expiration, explain gas exchange, and analyze respiratory disorders. Most of these structures are located inside the body and cannot be observed directly in classroom learning. Students must therefore construct mental representations from texts, two-dimensional images, videos, models, and verbal explanations. When the available representations are limited, students may memorize the names of organs without understanding the spatial and functional relationships among components. Virtual and augmented reality technologies can support the learning of anatomy and physiology, particularly when the content requires an understanding of bodily structures (Moro et al., 2021). The development of ARBOOK for anatomy learning has likewise shown that augmented reality representations can support three-dimensional interpretation and an understanding of spatial relationships, although the outcomes remain dependent on the design of the learning activities (Ferrer-Torregrosa et al., 2015).

Among lower secondary students, augmented reality has been combined with the structure-behavior-function model to connect the anatomy, function, and behavior of the circulatory system. This approach provides a relevant analogy for the respiratory system, which must also be understood as a network of interrelated components (Gnidovec et al., 2020). In anatomy education, mobile augmented reality has also been associated with higher achievement and lower cognitive load when abstract information is paired with focused visual representations (Küçük et al., 2016). These findings do not directly demonstrate the benefits of Assemblr EDU in the present study, but they provide a rationale for using representations that allow internal bodily structures to be examined from multiple perspectives.

This need was also evident in the learning context at UPT SPF SMP Negeri 33 Makassar. Preliminary observations and an interview with a science teacher indicated that instruction had incorporated LCD projectors, PowerPoint presentations, instructional videos, and digital modules. However, these media were still used primarily for one-way information delivery. The teacher also indicated that some students became bored easily, participated minimally during lessons, and had not consistently attained the expected mastery criterion in the human respiratory system unit. These preliminary observations provided the pedagogical context for selecting a medium that could offer three-dimensional representations and more concrete interaction with the subject matter.

Augmented reality (AR) integrates digital objects with the real environment through digital devices. AR enables students to examine three-dimensional models, rotate objects, enlarge specific components, and view relationships among structures from multiple angles. These characteristics may help students construct more concrete spatial representations. Wu et al. (2013) explained that AR can integrate digital information with the physical environment and provide multiple perspectives on learning objects. However, AR implementation also presents limitations, including device-related difficulties, application instability, connectivity constraints, and classroom management demands. Akçayır and Akçayır (2017) emphasized that the educational benefits of AR must be considered alongside the technical and pedagogical challenges that arise during implementation. At the adoption level, successful AR use is also influenced by teacher readiness, institutional support, device access, and the alignment between the technology and the learning environment; technological innovation therefore cannot be separated from the conditions under which it is implemented (De Lima et al., 2022).

AR can support learning outcomes, but the magnitude of its benefits is not consistent across contexts. Chang et al.'s (2022) meta-analysis of experimental and quasi-experimental studies identified benefits for learners' responses, knowledge, skills, and performance, while also showing that intervention duration was associated with variation in outcomes. At the secondary school level, Şahin and Yılmaz (2020) reported better achievement and attitudes in AR-supported science instruction than in conventional instruction. In biology learning, AR has also been used as a supplement that complements classroom content and activities rather than replacing the entire instructional process (Weng et al., 2020). Findings from other immersive technologies suggest potential support for STEM literacy, but virtual reality outcomes cannot be equated directly with augmented reality because the two technologies differ in their interaction characteristics and learning environments (Widiyatmoko et al., 2023). Nevertheless, these findings do not imply that AR automatically improves learning. Cuendet et al. (2013) showed that classroom AR design must account for the curriculum, teachers' needs, the flexibility of learning activities, and the realities of classroom implementation.

Studies with comparison groups can evaluate differences in achievement between two conditions, whereas a one-group design can show only that the second measurement differs from the first. These differences in design strength determine the language that can appropriately be used in drawing conclusions. In the present study, a higher posttest mean supports the statement that scores changed after instruction, but it does not establish that the change would not have occurred without Assemblr EDU. Distinguishing temporal sequence from causal attribution therefore underpins the formulation of the study's objectives, results, discussion, and conclusions.

Dual representational channels can help students organize information when verbal and visual sources complement one another. Three-dimensional models enable students to connect anatomical terms with the shapes and positions of organs, while teacher explanations help them interpret what they observe. However, presenting too many labels, movements, or information elements simultaneously may overload processing capacity. Research in physics laboratories has shown that AR can reduce extraneous cognitive load when real and virtual information are presented with appropriate spatial and temporal contiguity, although reduced cognitive load does not necessarily produce differences in conceptual achievement (Thees et al., 2020). Learning activities should therefore guide students' attention progressively toward relevant elements. Students may first be asked to identify the major organs and then examine relationships among components, respiratory mechanisms, and the application of concepts to health disorders. This sequencing is important if technology is to support the construction of mental models rather than merely provide an engaging visual experience. Because the present study did not measure these mental processes directly, the relationship between visualization and score changes is discussed only as a theoretically plausible explanation grounded in the literature.

Assemblr EDU is a learning platform that provides three-dimensional objects and AR features for educational activities (Assemblr, n.d.). In the human respiratory system unit, the platform can display organs, their relative positions, and structural relationships within the body. These capabilities are relevant to content that is difficult to observe directly. However, research on AR in science education has largely focused on total achievement, motivation, attitudes, and user experience (Arıcı et al., 2019).

The most closely related studies indicate that Assemblr EDU has already been used to teach the human respiratory system. Sekarani and Abdullah (2025) employed a quasi-experimental design with experimental and control classes among fifth-grade students and reported differences in learning outcomes following the use of Assemblr EDU. Ratih et al. (2025), meanwhile, developed Assemblr EDU-based augmented reality media integrated with contextual teaching and learning for the same topic and evaluated its feasibility and practicality. These studies demonstrate that the use of Assemblr EDU for respiratory-system instruction is no longer an entirely novel context. The contribution of the present study therefore lies not in the novelty of the application or content, but in the focused exploratory reporting of overall score changes, accompanied by supplementary C1-C4 item-group descriptions among eighth-grade students and interpretive limitations consistent with a one-group design.

Analysis by cognitive level requires caution because changes in the overall score do not directly identify the cognitive-process groups represented by the instrument. Classifying items into C1, C2, C3, and C4 can provide a more detailed description than a total score, but the strength of the interpretation depends on the number, quality, and difficulty of items in each group. In the present study, item allocation was unequal; consequently, results across levels were not treated as equivalent subscales. The analysis was used to describe observed supplementary response patterns among items classified as remembering, understanding, applying, and analyzing, rather than to identify the domain most strongly affected by the medium.

Rotation and magnification features may be useful when students need to examine the position or shape of an organ, but these features may be insufficient for explaining processes that change over time. Inspiration, expiration, and gas exchange require representations of movement, sequence, and pressure-volume relationships. If the displayed objects show anatomy only, teachers must supplement them with process explanations, analogies, flow diagrams, or cause-and-effect questions. Studies integrating AR with inquiry frameworks have demonstrated the importance of instructional models for directing observation and knowledge construction

(Wen et al., 2023), while AR integrated with problem-based learning has shown that digital representations acquire pedagogical value when used to solve structured problems (Fidan & Tuncel, 2019). Evaluating Assemblr EDU should therefore involve not only the presence of three-dimensional models but also the way those models are embedded in a sequence of learning activities. This principle also helps explain why AR studies can produce different outcomes even when similar technologies are used.

The relationship between three-dimensional representations and C1-C4 outcomes can be framed as a set of plausible pathways rather than as mechanisms that were directly tested. At C1, the shape and position of an organ may provide visual cues for recognizing terminology. At C2, spatial relationships among components may help students organize explanations of structure and function. At C3, representations may serve as references when concepts are applied to health situations or changes in organ function. At C4, connections among system components may support the analysis of cause-and-effect relationships. Research in geometry has shown that AR can support visualization skills when combined with structured thinking and problem-solving activities, making visualization part of a guided cognitive process rather than merely a display feature (Abdul Hanid et al., 2022). Nevertheless, each cognitive process requires different learning activities. Observing an object does not necessarily produce the ability to explain, apply, or analyze unless students are guided to process the information. Because the present study measured only test responses before and after instruction, attention, mental-model construction, and problem-solving strategies were not observed. This framework is therefore used to constrain the theoretical interpretation and to prevent score changes from being treated as direct evidence of specific cognitive mechanisms.

The study used a one-group pretest-posttest design without a control group. This design establishes only the sequence of score changes before and after instruction and cannot separate the contribution of Assemblr EDU from teacher explanations, practice, pretest experience, or other factors operating during the learning process. Accordingly, the terms change and descriptive improvement are used consistently, whereas claims of effect, effectiveness, and causality are avoided. This restriction is necessary to keep the conclusions aligned with the strength of the available evidence.

The study contributes overall score changes, with C1-C4 item-group results serving as supplementary descriptive evidence rather than comparable primary outcomes for the human respiratory system. Reporting by cognitive level may help teachers and researchers recognize that a total score is composed of items requiring different forms of thinking. However, this contribution remains exploratory because item allocation was unequal and detailed psychometric evidence was unavailable. This limited positioning is important for distinguishing findings directly supported by the data from conceptual explanations that require further testing through controlled designs and stronger instruments.

This study aimed to describe students' learning outcomes before and after Assemblr EDU-based augmented reality instruction on the human respiratory system, calculate relative change using N-gain based on group means, and provide supplementary change patterns among items classified as remembering, understanding, applying, and analyzing. The research questions were: (1) How did the mean learning-outcome score change from pretest to posttest? (2) What was the overall N-gain category? and (3) What descriptive patterns of change were observed in C1, C2, C3, and C4, given the unequal number of items? The study was positioned as a preliminary investigation of one class rather than as a basis for generalizing the effectiveness of Assemblr EDU to a broader population.

METHOD

This study employed a quantitative approach with a one-group pretest-posttest pre-experimental design. One group of students completed an initial test, received Assemblr EDU-based augmented reality instruction, and then completed a final test. The design was represented as O1-X-O2, where O1 denoted the pretest, X denoted instruction supported by Assemblr EDU, and O2 denoted the posttest. Because no comparison group was included, the analysis was intended to describe changes within the same group and was not used to infer comparative effectiveness or causality. The study involved eighth-grade students at UPT SPF SMP Negeri 33 Makassar. The population comprised 287 eighth-grade students. The sample consisted of 32 students from one class selected through purposive sampling. This technique produced a sample that was suitable for the study's implementation but did not give every member of the population an equal probability of selection. The findings therefore primarily describe the students in the participating class and cannot be generalized directly to all eighth-grade students in other schools.

The instructional variable was the use of Assemblr EDU-based augmented reality for the human respiratory system, and the outcome variable was students' test scores before and after instruction. Learning outcomes were operationalized through responses to multiple-choice items grouped according to the cognitive-process dimension of the revised taxonomy: remembering, understanding, applying, and analyzing (Anderson & Krathwohl, 2001). The test content covered the structure and function of respiratory organs, the air pathway, the mechanisms of inspiration and expiration, gas exchange, and respiratory-system disorders. The procedure comprised three main stages. First, a pretest was administered to the 32 students to obtain their initial learning-outcome scores. Second, the human respiratory system was taught using Assemblr EDU-based augmented reality media. The medium presented three-dimensional visual representations related to the instructional content. Third, the same group completed a posttest after instruction. The pretest and posttest assessed the same cognitive indicators, allowing changes in mean scores to be described.

Assemblr EDU was positioned as a visual-representation medium within the instructional sequence rather than as the sole learning resource. The three-dimensional objects enabled the respiratory-system content to be presented in ways that could help students identify shapes and relationships among components. However, the study did not report the number of sessions, instructional duration, device configuration, or the complete sequence of classroom activities. Consequently, no claims were made about any specific intervention component, and the interpretation was limited to score changes following instruction that involved the medium. The research instrument was a 20-item multiple-choice achievement test. Each item had one correct answer and several distractors. Correct answers were scored as one and incorrect answers as zero, after which raw scores were converted to a 0-100 scale. The instrument comprised five C1 items, seven C2 items, two C3 items, and six C4 items. This unequal allocation meant that the scores for the cognitive groups differed in stability. Accordingly, total scores remained the primary outcome; C1-C4 scores provided supplementary descriptions rather than four fully comparable subscales.

The instrument was reported to have undergone expert validation for content alignment, indicators, item construction, and language. However, the number and expertise of the validators, the content-validity coefficient, reliability results, item difficulty, discrimination, and distractor functioning were not reported. This limited psychometric information must be considered when interpreting the findings. The most evident problem concerned C3, which was represented by only two items; a single correct or incorrect response could therefore contribute substantially to the percentage score for that group. A two-item measure also requires a specific approach to reliability estimation and cannot be treated in the same manner as a subscale with a larger number of items (Eisinga et al., 2013). The C3 result was therefore not used to conclude that the instruction most strongly supported applying.

The scores produced by the instrument represented an operational measure of learning outcomes based on the available items rather than a comprehensive representation of all competencies related to the respiratory system. Multiple-choice tests can cover content relatively broadly and permit consistent scoring, but the quality of interpretation still depends on the alignment of the stimulus, response options, and cognitive demands of each item. Classifying an item as C1, C2, C3, or C4 cannot be based solely on the verb used in the indicator; the actual thinking required to obtain the answer must also be considered. An item framed as analysis may still be answered through recall when the options are overly obvious. Conversely, a brief item may require reasoning if students must connect several pieces of information. Because representative items and empirical item analyses were not presented, the cognitive classifications were treated as the classifications used by the researchers rather than as psychometric evidence that four distinct constructs had been measured.

Descriptive analysis was used to compare the pretest and posttest means. Relative change was calculated using normalized gain based on group means: g = (posttest mean - pretest mean)/(maximum score - pretest mean). The maximum score was 100. Classification followed Hake (1998), values below 0.30 were categorized as low, values from 0.30 to below 0.70 as moderate, and values of 0.70 or above as high. The calculation was performed for the overall score and for each cognitive item group. The analysis did not include standard deviations, medians, ranges, confidence intervals, paired tests, p values, or effect sizes because these statistics were unavailable. The findings were therefore not described as statistically significant. Means and N-gain values were used to characterize the direction and magnitude of descriptive change in the participating group. This limitation is important because an N-gain category cannot replace inferential testing and does not demonstrate that the medium caused the score changes.

RESULTS AND DISCUSSION

The analysis showed that the mean learning-outcome score changed from 42.03 on the pretest to 82.81 on the posttest. The mean difference between the two measurements was 40.78 points. Applying the formula to the group means produced an overall N-gain of 0.70, which was classified as high. Descriptively, the participating group achieved a higher mean score after instruction involving Assemblr EDU-based augmented reality (see Table 1).

Table 1. Changes in learning-outcome scores before and after instruction

Measurement	Number of students	Mean score
Pretest	32	42.03
Posttest	32	82.81
Mean difference		40.78
N-gain		0.70
N-gain category		High

The N-gain in Table 1 was calculated from the group means. A mean score summarizes change at the class level but does not describe variation among students. Without standard deviations, ranges, and individual-level data, it is not possible to determine whether the change was relatively uniform or was driven primarily by a subset of students. Interpretation is therefore limited to the difference between group means and the N-gain category, rather than statistical significance or the consistency of change across participants.

N-gain provides information that differs from the raw mean difference. The 40.78-point difference represents the distance between the pretest and posttest means, whereas an N-gain of 0.70 expresses the change relative to the remaining possible improvement from the initial score to the maximum score. Although useful as a summary measure, N-gain is influenced by the initial score. Groups with higher initial scores have less room for improvement, while groups with lower initial scores have more. N-gain values across item groups therefore cannot be compared directly as measures of the strength of the medium. Differences in baseline scores, numbers of items, and possible item difficulty also contribute to the resulting values. In this study, N-gain addressed a descriptive question concerning relative change; it was not used to test hypotheses, determine statistical significance, or estimate the causal effect of Assemblr EDU.

The respiratory system comprises internal organs, spatial relationships, and processes that are difficult to observe directly. Three-dimensional models can provide visual references that allow students to identify the position and arrangement of structures more concretely than text alone. However, the medium was not the only component of instruction. Teacher explanations, exposure to the content, practice, and experience with the test could also have contributed to posttest scores. The findings are therefore more appropriately understood as changes following the overall instructional experience involving Assemblr EDU, rather than as an independent effect of the technology.

This pattern is broadly consistent with previous evidence. Garzón and Acevedo (2019) found that AR was associated with moderate average learning gains, although the magnitude varied according to learner characteristics, subject area, learning environment, and comparison condition. Wahyu et al. (2020) likewise reported differences in science achievement and scientific literacy between mobile AR-assisted STEM instruction and conventional instruction. However, evidence on AR does not consistently demonstrate an independent technological advantage. Faria (2024) showed that teaching strategy and its combination with AR can produce different patterns of achievement and retention, indicating that technology should be analyzed together with the instructional design in which it is embedded. Comparisons with these studies must remain cautious because the present study used no control group and described change within a single class only.

Pedagogical approaches also influence how technology functions in learning. Garzón et al. (2020) showed that instructional approaches moderate the impact of AR, while Setiawaty et al. (2024) demonstrated that integrating mobile AR with inquiry activities can support science-process skills and conceptual mastery. These findings provide a basis for interpreting three-dimensional models as resources that must be linked to relevant explanations, questions, and tasks. Because the present study did not compare instructional components separately, however, the contributions of visualization, teacher explanation, practice, and test experience could not be isolated.

Distinguishing the medium from the learning activities is important when interpreting the findings. Assemblr EDU provides a means of displaying objects, but the construction of understanding still depends on what students do with those objects. Two classes may use the same application yet experience different learning opportunities if one merely views the display while the other connects it to questions, explanations, and problem solving. The present study included no comparison condition that could separate these possibilities. The term Assemblr EDU-supported instruction is therefore more appropriate than treating the medium as a single, self-contained intervention. This position also explains why the practical implications do not advocate automatic adoption of the application, but instead emphasize alignment among visual representation, cognitive objectives, content, and assessment.

A one-group design requires cautious interpretation. Score changes may be associated with the use of the medium, teacher presentation, practice during instruction, experience with the pretest, or other factors occurring between the two measurements. Without a group learning the same content under a different condition, the changes observed in the participating class cannot be compared with changes that might have occurred without Assemblr EDU. The study therefore does not support the claim that Assemblr EDU caused the increase or was more effective than other media.

In addition to overall scores, supplementary patterns were described by cognitive-process group. C1 changed from 55.6 to 83.1, with an N-gain of 0.62; C2 changed from 41.0 to 79.4, with an N-gain of 0.65; C3 changed from 33.0 to 87.5, with an N-gain of 0.81; and C4 changed from 42.7 to 84.8, with an N-gain of 0.73. All item groups had higher posttest scores. Direct comparisons, however, must consider the unequal numbers of items and the absence of reliability and difficulty information for each group (see Table 2).

Table 2. Supplementary score changes across cognitive-process groups

Level	Items	Pretest	Posttest	N-gain	Category
C1-Remembering	5	55.6	83.1	0.62	Moderate
C2-Understanding	7	41.0	79.4	0.65	Moderate
C3-Applying	2	33.0	87.5	0.81	High
C4-Analyzing	6	42.7	84.8	0.73	High

Remembering, or C1, involves recognizing and retrieving previously learned information. In the respiratory-system unit, this ability includes recalling organ names, basic terminology, the air pathway, and simple functions. The C1 pretest score of 55.6 was the highest initial score among the four item groups and increased to 83.1, yielding a moderate N-gain of 0.62. Conceptually, organ representations may provide visual cues regarding shape and position that help students connect terminology with the objects they observe. However, the finding reflects performance on only five multiple-choice items classified as C1. The multiple-choice format should be considered when interpreting remembering. Students may select a correct answer by recognizing a term or an available option, whereas unaided retrieval may produce different performance. Multiple-choice tests are efficient and permit objective scoring, but they also allow correct responses through recognition or guessing. The C1 score therefore does not necessarily indicate deep factual mastery. Future research could combine multiple-choice items with image labeling, short-answer questions, and delayed tests to assess both recall accuracy and retention.

The lower C1 N-gain relative to C3 and C4 does not mean that the medium was less supportive of remembering. The C1 baseline was already higher, leaving less room for relative improvement. N-gain is also influenced by the initial score and item characteristics. Because the study reported only immediate pretest and posttest measurements and included no delayed test, the C1 result is more appropriately interpreted as a short-term change in performance on the instrument rather than evidence of long-term retention. Understanding, or C2, includes explaining, interpreting, comparing, classifying, and inferring. In the respiratory-system unit, students need to understand organ functions, structure-function relationships, the air pathway, and the differences between inspiration and expiration. The C2 score changed from 41.0 to 79.4, yielding a moderate N-gain of 0.65. The change indicates that students answered more of the seven understanding items correctly after instruction. Three-dimensional representations may help organize information about relationships among the trachea, bronchi, bronchioles, and lungs, but the test results do not show whether students could independently explain their reasoning.

Physiological processes cannot be understood fully from anatomical form alone. Gas exchange, changes in thoracic volume, and the relationship between pressure and airflow require explanations of sequence and mechanism. Ibáñez and Delgado-Kloos (2018) emphasized the importance of instructional and metacognitive support in AR-based STEM learning. Evidence from digital materials with three-dimensional displays likewise suggests that visualization becomes more meaningful when linked to problem-solving activities and the development of conceptual understanding (Distrik et al., 2024). The C2 score change may therefore reflect a combination of visual representation, content presentation, and learning activities. Because the contribution of each component was not measured separately, the medium cannot be treated as the sole explanation for the change.

Applying, or C3, involves using concepts or procedures in particular situations. In the respiratory-system unit, this may include applying concepts to explain changes in breathing rate, predict the consequences of organ dysfunction, or select actions based on respiratory-health principles. The C3 score changed from 33.0 to 87.5, yielding a high N-gain of 0.81. Numerically, this group showed the largest relative change, but C3 was measured with only two items, whereas the other groups contained five to seven items. Two items are insufficient to form a stable measure of applying. A single correct or incorrect response contributes substantially to the group percentage, and the lowest C3 baseline also provided the greatest room for improvement. The value of 0.81 therefore cannot support the conclusion that Assemblr EDU most strongly improved applying. The defensible statement is that the proportion of correct responses on the two items classified as C3 was higher on the posttest. Future studies require more applying items across varied contexts and adequate psychometric analysis.

Analyzing, or C4, requires students to break information into parts, identify relationships among components, distinguish relevant information, and explain cause-and-effect relationships (Anderson & Krathwohl, 2001). In the respiratory-system unit, these demands may involve examining the consequences of alveolar damage, bronchiolar narrowing, pollution exposure, or impaired gas exchange. The C4 score changed from 42.7 to 84.8, yielding a high N-gain of 0.73. Six items provided broader coverage than the two C3 items, but reliability, difficulty, and discrimination evidence for the C4 group was not reported, so interpretation remains limited.

Three-dimensional models may conceptually support analysis because the system is represented as a set of interrelated components (Garzón et al., 2020; Setiawaty et al., 2024). Students can use structural representations as references when considering functional relationships and the consequences of changes to components. Annisa and Subiantoro (2022) reported that mobile AR in respiratory-system instruction could support analytical ability and socioscientific reasoning, although overall conceptual outcomes did not always differ. Their findings provide a theoretical context for the C4 score, but differences in design and instrumentation preclude direct equivalence with the present study.

As supplementary evidence, the C1-C4 pattern shows that posttest scores were higher on factual items and on different items classified as understanding, applying, and analyzing. This pattern should be treated descriptively rather than as a hierarchy of the cognitive processes most strongly affected by the medium. Unequal item allocation, possible differences in item difficulty, and the absence of reliability evidence for each group reduce the comparability of the results. The most defensible conclusion is that the proportion of correct responses increased across all item groups used in the study.

Changes across all item groups may also indicate that the learning experience was not limited to recognizing organ names. However, this breadth cannot be regarded as evidence of transfer across cognitive levels. Students who answered C3 or C4 items correctly may not consistently apply or analyze concepts in other contexts. Demonstrating transfer would require multiple tasks in each domain, varied stimuli, and possibly open-ended items that reveal students' reasoning. With the available data, interpretation is limited to performance on the 20 multiple-choice items. This restriction does not eliminate the descriptive value of the findings but places them in proportion: the scores offer preliminary information about changes in test responses, whereas conclusions about the development of thinking skills require broader measurement evidence.

The score changes also cannot be used to conclude that motivation, attention, or engagement increased, or that cognitive load decreased, because these variables were not measured. The literature suggests that AR can influence attention, relevance, confidence, and learner satisfaction (Prasetya et al., 2024), but these mechanisms remain conceptual possibilities in the present study. Further investigation would require validated questionnaires, systematic observation, interviews, or interaction data to examine the relationship between the experience of using the medium and learning outcomes. AR implementation may also be influenced by instructional duration, device availability, technical capacity, and the way the medium is integrated into classroom activities. Chang et al. (2022) showed that intervention duration is associated with variation in AR research outcomes. However, the present study did not report the number and duration of sessions in detail or measure users' technical experiences. The findings therefore provide no evidence regarding time efficiency, ease of use, device readiness, or the sustainability of Assemblr EDU implementation in schools.

A limited practical implication is that Assemblr EDU may be considered a representational resource for topics requiring visualization of internal structures. Its use should be guided by clear observation objectives and linked to conceptual explanations. Teachers may use three-dimensional models as references when discussing organ names, positions, functions, the air pathway, and relationships among components. This recommendation, however, is based on the characteristics of the medium and descriptive change in one class, not on evidence that Assemblr EDU is more effective than other media.

In instructional planning, three-dimensional representations should be selected according to the conceptual problem they are intended to address. Visualizing organ positions is relevant to spatial relationships, whereas inspiration, expiration, and gas exchange also require explanations of sequence, movement, and pressure-volume relationships. Teachers should therefore distinguish between aspects of the content that can be clarified through three-dimensional form and those requiring process diagrams, animation, analogy, or demonstration. This principle follows from the characteristics of the content and medium rather than from comparative testing in the present study. Assemblr EDU may be positioned as one element within a combination of learning resources. Its evaluation should consider not only final scores but also conceptual accuracy, misconceptions, explanatory ability, and the persistence of understanding after instruction.

Assessment should align with the learning objectives and provide adequate item coverage. When the objectives include applying and analyzing, the instrument should measure these abilities across several contexts rather than through only one or two questions. The two C3 items in this study were insufficient to form a stable subscale. Future research could use cases involving physical activity, air pollution, smoking, respiratory disease, and organ damage to broaden the range of application. Each item should also be supported by evidence of content validity, difficulty, discrimination, distractor functioning, and reliability.

The principal limitations were a one-group design without a comparison condition and a narrow focus on only a single instructional context. Nevertheless, transparent reporting retains methodological value. The available data show how conclusions must change when claims are aligned with the design and the quality of measurement. Means, score differences, and N-gain provide preliminary information that can inform future research design, identify the need for more balanced instruments, and indicate implementation features that should be documented. These preliminary findings should not, however, serve as the sole basis for large-scale adoption decisions. Replication with documented procedures, comparison groups, and individual-level analysis is required to test whether similar patterns occur across classes, schools, and student populations.

CONCLUSIONS

This study found that middle school students demonstrated higher learning outcomes after participating in Assemblr EDU-based augmented reality instruction on the human respiratory system. Improvement was observed not only in overall achievement but also across items representing remembering, understanding, applying, and analyzing cognitive processes. These findings suggest that Assemblr EDU can function as a useful three-dimensional representational resource for supporting the learning of anatomical structures and relationships that are difficult to observe directly. Its instructional value, however, depends on how the visual representations are integrated with teacher explanations, conceptual activities, and appropriately aligned assessment tasks.

The findings should be interpreted cautiously because the study employed a one-group pretest–posttest design without a comparison group and relied primarily on descriptive analysis. Therefore, the observed learning changes cannot be attributed solely to Assemblr EDU or interpreted as evidence of causal effectiveness. Future research should employ controlled or quasi-experimental designs, involve broader and more diverse samples, use more balanced and psychometrically validated assessment instruments, and examine additional outcomes such as conceptual understanding, retention, engagement, and cognitive processes. Such studies would provide stronger evidence regarding the educational contribution of augmented reality in science learning.

ACKNOWLEDGMENT

The authors would like to express their sincere appreciation to UPT SPF SMP Negeri 33 Makassar for providing permission and support during the implementation of this study. The authors also extend their gratitude to the science teacher and the participating students for their cooperation throughout the data collection and learning activities. Appreciation is also conveyed to the academic community of Universitas Negeri Makassar for the academic support provided during the preparation of this research.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm their contributions to this study as follows: MH contributed to the conceptualization, research design, data collection, data analysis, interpretation of the findings, and preparation of the initial manuscript. NHM contributed to the research methodology, supervision, interpretation of the findings, manuscript review, and critical revision. RNA contributed to the validation of the research process, interpretation of the findings, manuscript review, and revision of the article. All authors discussed the results, reviewed the manuscript, and approved the final version for publication.

AI DISCLOSURE STATEMENT

The authors declare that artificial intelligence-assisted tools were used solely to support language refinement, grammatical correction, and improvement of manuscript readability and organization. These tools were not used to generate research data, conduct statistical analysis, determine research findings, or make substantive scientific interpretations. All AI-assisted outputs were critically reviewed, verified, and revised by the authors, who remain fully responsible for the accuracy, originality, integrity, and final content of the manuscript.

REFERENCES

Abdul Hanid, M. F., Mohamad Said, M. N. H., Yahaya, N., & Abdullah, Z. (2022). Effects of augmented reality application integration with computational thinking in geometry topics. *Education and Information Technologies, 27*, 9485–9521. https://doi.org/10.1007/s10639-022-10994-w

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

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Annisa, D. N., & Subiantoro, A. W. (2022). Mobile augmented reality in socioscientific issues-based learning: The effectiveness on students' conceptual knowledge and socioscientific reasoning. *Jurnal Pendidikan IPA Indonesia, 11*(4), 611–625. https://doi.org/10.15294/jpii.v11i4.38993

Arıcı, F., Yıldırım, P., Çalıklar, Ş., & Yılmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education, 142*, 103647. https://doi.org/10.1016/j.compedu.2019.103647

Assemblr. (n.d.). *Assemblr EDU: Immersive 3D/AR platform for teachers & learning*. Retrieved July 11, 2026, from https://edu.assemblrworld.com/

Chang, H.-Y., Binali, T., Liang, J.-C., Chiou, G.-L., Cheng, K.-H., Lee, S. W.-Y., & Tsai, C.-C. (2022). Ten years of augmented reality in education: A meta-analysis of (quasi-)experimental studies to investigate the impact. *Computers & Education, 191*, 104641. https://doi.org/10.1016/j.compedu.2022.104641

Cuendet, S., Bonnard, Q., Do-Lenh, S., & Dillenbourg, P. (2013). Designing augmented reality for the classroom. *Computers & Education, 68*, 557–569. https://doi.org/10.1016/j.compedu.2013.02.015

De Lima, C. B., Walton, S., & Owen, T. (2022). A critical outlook at augmented reality and its adoption in education. *Computers and Education Open, 3*, 100103. https://doi.org/10.1016/j.caeo.2022.100103

Distrik, I. W., Ertikanto, C., Purwati, Y. S., Saregar, A., & Ab Rahman, N. F. (2024). Digital problem-based worksheet with 3D PageFlip: An effort to address concept understanding problems and enhance digital literacy skills. *Jurnal Pendidikan IPA Indonesia, 13*(1), 116–127. https://doi.org/10.15294/jpii.v13i1.48604

Eisinga, R., te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman–Brown? *International Journal of Public Health, 58*(4), 637–642. https://doi.org/10.1007/s00038-012-0416-3

Faria, A. (2024). Augmented reality and teaching strategies in the study of volcanism in elementary and secondary schools. *Journal of New Approaches in Educational Research, 13*, 18. https://doi.org/10.1007/s44322-024-00018-5

Ferrer-Torregrosa, J., Torralba, J., Jimenez, M. A., García, S., & Barcia, J. M. (2015). ARBOOK: Development and assessment of a tool based on augmented reality for anatomy. *Journal of Science Education and Technology, 24*, 119–124. https://doi.org/10.1007/s10956-014-9526-4

Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem based learning: The effects on learning achievement and attitude in physics education. *Computers & Education, 142*, Article 103635. https://doi.org/10.1016/j.compedu.2019.103635

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review, 27*, 244–260. https://doi.org/10.1016/j.edurev.2019.04.001

Garzón, J., Kinshuk, Baldiris, S., Gutiérrez, J., & Pavón, J. (2020). How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. *Educational Research Review, 31*, 100334. https://doi.org/10.1016/j.edurev.2020.100334

Gnidovec, T., Žemlja, M., Dolenec, A., & Torkar, G. (2020). Using augmented reality and the structure–behavior–function model to teach lower secondary school students about the human circulatory system. *Journal of Science Education and Technology, 29*, 774–784. https://doi.org/10.1007/s10956-020-09850-8

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

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

Küçük, S., Kapakin, S., & Göktaş, Y. (2016). Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. *Anatomical Sciences Education, 9*(5), 411–421. https://doi.org/10.1002/ase.1603

Moro, C., Birt, J., Stromberga, Z., Phelps, C., & Clark, J. (2021). Virtual and augmented reality enhancements to medical and science student physiology and anatomy test performance: A systematic review and meta-analysis. *Anatomical Sciences Education, 14*(3), 368–376. https://doi.org/10.1002/ase.2049

Prasetya, F., Fortuna, A., Samala, A. D., Rawas, S., Mystakidis, S., Syahril, Waskito, Primawati, Wulansari, R. E., & Kassymova, G. K. (2024). The impact of augmented reality learning experiences based on the motivational design model: A meta-analysis. *Social Sciences & Humanities Open, 10*, 100926. https://doi.org/10.1016/j.ssaho.2024.100926

Ratih, C. D., Anggoro, S., & Fukui, M. (2025). Development of augmented reality-based learning media using Assemblr EDU integrated with contextual teaching and learning (CTL) for the human respiratory system. *Jurnal Penelitian Pendidikan IPA, 11*(8), 720–732. https://doi.org/10.29303/jppipa.v11i8.12116

Şahin, D., & Yılmaz, R. M. (2020). The effect of augmented reality technology on middle school students' achievements and attitudes towards science education. *Computers & Education, 144*, 103710. https://doi.org/10.1016/j.compedu.2019.103710

Sekarani, F., & Abdullah, K. (2025). Effect of Assemblr Edu on fifth graders' learning outcomes in the respiratory system. *Inovasi Kurikulum, 22*(3), 1635–1646. https://doi.org/10.64014/jik.v22i3.91

Setiawaty, S., Lukman, I. R., Imanda, R., Sudirman, S., & Rauzatuzzikrina, R. (2024). Integrating of mobile augmented reality applications through inquiry learning to improve students' science process skills and concept mastery. *Jurnal Pendidikan IPA Indonesia, 13*(1), 90–102. https://doi.org/10.15294/jpii.v13i1.48891

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior, 108*, 106316. https://doi.org/10.1016/j.chb.2020.106316

Wahyu, Y., Suastra, I. W., Sadia, I. W., & Suarni, N. K. (2020). The effectiveness of mobile augmented reality assisted STEM-based learning on scientific literacy and students' achievement. *International Journal of Instruction, 13*(3), 343–356. https://doi.org/10.29333/iji.2020.13324a

Wen, Y., Wu, L., He, S., Ng, N. H.-E., Teo, B. C., Looi, C. K., & Cai, Y. (2023). Integrating augmented reality into inquiry-based learning approach in primary science classrooms. *Educational Technology Research and Development, 71*, 1631–1651. https://doi.org/10.1007/s11423-023-10235-y

Weng, C., Otanga, S., Christianto, S. M., & Chu, R. J.-C. (2020). Enhancing students' biology learning by using augmented reality as a learning supplement. *Journal of Educational Computing Research, 58*(4), 747–770. https://doi.org/10.1177/0735633119884213

Widiyatmoko, A., Nugrahani, R., Yanitama, A., & Darmawan, M. S. (2023). The effect of virtual reality game-based learning to enhance STEM literacy in energy concepts. *Jurnal Pendidikan IPA Indonesia, 12*(4), 648–657. https://doi.org/10.15294/jpii.v12i4.48265

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024