Lensa: Jurnal Kependidikan Fisika
Penerbit: Universitas Pendidikan Mandalika
https://e-journal.undikma.ac.id/index.php/Lensa/index

Dec. 2026, Vol. 14, No. 2
e-ISSN: 2686-0937
pp. 179–204

Sound Wave Conceptual Understanding through Experiment-Based and Multimodal Meaningful Instructional Design

Ansori*, Ahmad Harjono, Ni Nyoman Sri Putu Verawati
Physics Education Department, Universitas Mataram, Mataram, Indonesia
Corresponding Author: ansori041203@gmail.com

Abstract

Understanding sound waves requires students to connect observable vibrations with abstract concepts of propagation, medium interactions, and wave behavior. This study examined conceptual understanding of sound waves through experiment-based and multimodal *Meaningful Instructional Design* (MID) compared with conventional instruction. A quasi-experimental nonequivalent control group design involved 52 eleventh-grade students from two intact classes at an Indonesian senior high school, with 26 students in each group. Over three instructional meetings, the experimental group engaged in MID activities organized into *lead-in*, *reconstruction*, and *production* stages, integrating hands-on experiments, worksheets, auditory and visual representations, and digital simulations. Conceptual understanding was assessed using seven constructed-response items administered before and after instruction. Data were analyzed using descriptive statistics, independent-samples t-tests, normalized gain, and analysis of covariance (ANCOVA) controlling for pretest scores. Mean scores increased from 44.15 to 88.15 in the experimental group and from 53.88 to 83.27 in the control group. The unadjusted posttest difference was not statistically significant (p = .077), whereas score gains differed significantly (p = .007). ANCOVA indicated a significant adjusted group difference, $F(1, 49) = 7.58$, p = .008, partial eta squared (ηp^2) = .134, with adjusted means of 89.16 and 82.26, respectively. Mean normalized gains were .778 and .632. The analysis also characterized changes across six cognitive process indicators. All six cognitive indicators improved, with the largest between-group difference in change occurring for analyzing (C4). These findings suggest that integrated MID instruction is associated with stronger conceptual understanding after accounting for initial achievement, supporting its consideration in sound wave instruction.

Keywords: Conceptual understanding; Sound waves; Meaningful instructional design; Multimodal learning; Physics experiments

How to cite: Ansori, A., Harjono, A., & Verawati, N. N. S. P. (2026). Sound Wave Conceptual Understanding through Experiment-Based and Multimodal Meaningful Instructional Design. *Lensa: Jurnal Kependidikan Fisika*, *14*(2), 179–204. https://doi.org/10.33394/j-lkf.v14i2.22525

INTRODUCTION

Conceptual understanding is a fundamental objective of physics education because it enables students to explain natural phenomena through relationships among scientific concepts, principles, and representations. Such understanding extends beyond recalling definitions or applying mathematical equations to include interpreting phenomena, integrating information, and transferring physical principles to unfamiliar contexts. Sound waves are particularly demanding because students must relate source vibrations, medium properties, energy transmission, and acoustic phenomena. Research has shown that students' reasoning about sound propagation frequently diverges from accepted physical models (Wittmann et al., 2003). These difficulties are associated with the mental models students construct and the connections they make between everyday experiences and conceptual explanations (Hrepic et al., 2010). The development of diagnostic instruments for sound concepts further suggests that conceptual understanding

cannot be captured adequately by knowledge of terminology or mathematical procedures alone (Eshach, 2014).

A persistent challenge is that students often conceive of sound as a substance or object traveling from a source to a receiver. This conception conflicts with the scientific account of sound as a mechanical disturbance transmitted through a medium. Cross-sectional evidence indicates that materialistic reasoning about sound can persist even after formal instruction in acoustics (Eshach et al., 2018). Similar difficulties arise when students explain simple acoustic devices, a task that requires coordination of physical observations with the characteristics of wave propagation (Volfson et al., 2018). Secondary school students may simultaneously accept scientific explanations and retain materialistic ideas about sound (Eshach et al., 2016). Accordingly, instruction must address how students categorize sound as a physical process rather than simply emphasize memorization. From a conceptual-change perspective, these difficulties relate to ontological distinctions between objects and processes (Chi et al., 2012; Volfson et al., 2019).

The complexity of sound-wave understanding becomes more apparent when fundamental concepts are applied to phenomena involving several physical variables. Standing waves in air columns, resonance, interference, and frequency shifts require students to coordinate visual information, mathematical relationships, and mechanistic explanations. Analyses of textbook illustrations of standing waves in pipes indicate that representations require careful design because images influence how students interpret particle motion and wave patterns (Zeng et al., 2014). Historical accounts of sound concepts likewise underscore the distinction between substance-based and wave-process explanations (Volfson et al., 2022). More broadly, conceptual surveys of mechanical waves indicate that propagation, superposition, reflection, and standing waves involve interrelated knowledge components (Barniol & Zavala, 2016). Students' consistency in applying such concepts may also vary with the problem context (Tongchai et al., 2011).

Conceptual difficulties do not necessarily imply an absence of relevant prior knowledge. The conceptual-resources perspective recognizes that everyday experiences can generate intuitive ideas that may support scientific understanding. Research on mechanical-wave propagation found that students draw on diverse reasoning resources, including intuitive understandings of media, pulse motion, and energy (Goodhew et al., 2019). The use of these resources is context-sensitive and may shift across question formats (Goodhew et al., 2021). Conceptually oriented instruction should therefore create opportunities to activate initial ideas, compare them with physical evidence, and develop more coherent conceptual relationships. Research on knowledge integration in mechanical-wave propagation emphasizes connecting conceptual components previously understood in isolation to produce a more integrated knowledge structure (Xie et al., 2021).

Conceptual development also depends on students' ability to use different representations of physical phenomena. Sound waves may be expressed through source motion, pressure variations, wave patterns, graphs, equations, and auditory experiences. Each representation captures particular features of the same phenomenon; comprehensive understanding therefore requires coordination across formats. Research on analogies in physics learning indicates that representations shape how students relate phenomena to explanatory models

(Podolefsky & Finkelstein, 2006). The ability to select and evaluate representations is also associated with physics problem solving, especially when information must be translated between formats (Kohl & Finkelstein, 2005). Sound-wave instruction consequently needs to help students connect observable phenomena to the conceptual representations used to explain them.

A multimodal approach offers opportunities to present physics concepts through verbal, visual, auditory, and empirical experiences. Nevertheless, providing multiple representations does not automatically improve conceptual understanding. The Design, Functions, Tasks (DeFT) framework emphasizes that the effectiveness of multiple representations depends on their respective functions, design characteristics, and the cognitive activities they elicit (Ainsworth, 2006). Multimedia learning principles likewise stress the integration of verbal and visual information rather than the mere addition of media (Mayer, 2003). When information is presented in several formats, learners must recognize meaningful connections among them to construct coherent knowledge (Seufert, 2003). This principle is particularly relevant to sound waves because audible and observable phenomena must be related to propagation mechanisms that are not directly visible.

Experiments can connect physics concepts to empirical phenomena. Observation and variable manipulation allow students to identify regularities, evaluate explanations, and construct evidence-based understanding. However, some features of sound waves, such as pressure variations and medium-particle dynamics, cannot readily be observed directly. Digital representations and virtual simulations may help visualize these features. A systematic review of combined physical and virtual experiments found that these environments can provide complementary conceptual support, although no implementation sequence is uniformly superior across learning contexts (Wörner et al., 2022). A review of Physics Education Technology (PhET) simulations likewise highlights their relevance to conceptual physics learning when embedded in guided activities (Banda & Nzabahimana, 2021). More recent meta-analytic evidence indicates that combining real and virtual experiments can support physics achievement, with effectiveness varying according to content, learner characteristics, and instructional design (Wang et al., 2025).

Experiments and multimodal representations also require instructional organization that supports active construction of scientific explanations. Observing a phenomenon alone does not guarantee conceptual understanding unless students can relate what they observe to relevant physical principles. A meta-analysis of inquiry-based science teaching identified implementation features and instructional guidance as important determinants of learning outcomes (Furtak et al., 2012). Another meta-analysis likewise reported the importance of guidance in inquiry-based learning (Lazonder & Harmsen, 2016). Together, these studies suggest that experiments should be accompanied by conceptual prompts, meaningful questions, reflection, and opportunities to apply knowledge. Instruction integrating firsthand experience and multiple representations therefore benefits from a sequence that connects prior knowledge, new evidence, and conceptual application.

Meaningful Instructional Design (MID) offers one framework for organizing

these connections. In this study, MID refers to instruction that emphasizes relationships between students' prior knowledge, meaningful learning experiences, and the application of knowledge in relevant situations. The framework includes activating prior knowledge during lead-in, developing conceptual relationships during reconstruction, and applying and communicating understanding during production. These stages organize learning experiences around the construction of meaning. In sound-wave instruction, integrating MID with experiments and multimodal representations is relevant because it provides opportunities to connect vibrations, medium properties, wave representations, and acoustic applications. The relationships among these instructional components provide a conceptual basis for learning experiences that extend beyond information delivery to the development of interconnected understanding.

Physics education research has documented difficulties in understanding sound waves, the importance of knowledge integration, the value of multiple representations, and the potential of combining real and virtual experiments. However, this literature has primarily addressed conceptual difficulties or individual instructional components. Studies of sound tend to focus on misconceptions and mental models, whereas multimedia and experimental studies examine how different information sources and learning experiences support understanding. Bringing these strands together creates a rationale for examining MID, experimentation, and multimodal representations within an integrated framework for teaching sound waves. This focus is pertinent because conceptual understanding involves not only achievement gains but also students' ability to relate empirical phenomena to scientific explanations across cognitive processes.

Research Objectives and Questions

This study aimed to examine students' conceptual understanding of sound waves under experiment-based and multimodal MID instruction compared with conventional instruction. Specifically, it addressed differences in conceptual understanding between instructional conditions, changes from pretest to posttest, and patterns of change across six cognitive-process indicators. These objectives focus on the relationship between the organization of learning experiences and sound-wave conceptual understanding in secondary education.

The study addressed the following research questions:

1. How does sound-wave conceptual understanding differ between students receiving experiment-based and multimodal MID and those receiving conventional instruction after accounting for initial achievement?
2. How do absolute and relative changes in sound-wave conceptual understanding compare between the two instructional groups?
3. What patterns of change are observed across the cognitive indicators C1 (remembering), C2 (understanding), C3 (applying), C4 (analyzing), C5 (evaluating), and C6 (creating) in the two groups?

Research Novelty

The distinctive contribution of this study is its examination of sound-wave conceptual understanding within a framework that integrates MID, hands-on experimentation, and multimodal representations as a coherent learning experience. Whereas previous studies have often investigated sound misconceptions, representational approaches, or real and virtual experiments

separately, the present study focuses on instruction linking prior-knowledge activation, conceptual reconstruction, and knowledge application. Its intended contribution extends beyond comparing achievement across instructional conditions to characterizing changes across cognitive indicators C1-C6. This focus adds empirical evidence on sound-wave teaching that combines firsthand experience, multiple representations, and meaning-oriented instructional organization.

METHODS

Research Design and Participants

The study employed a quantitative quasi-experimental design with nonequivalent control groups. Two intact classes completed a pretest, received instruction under their assigned conditions, and completed a posttest; students were not randomly assigned. The design compared sound-wave conceptual understanding among students taught through experiment-based and multimodal Meaningful Instructional Design (MID) and those receiving conventional instruction.

The study was conducted in May 2026 at SMAN 2 Labuapi, West Lombok Regency, West Nusa Tenggara, Indonesia. The accessible population comprised 52 eleventh-grade students in two classes. Total population sampling was used to include all students in these classes. Class XI A2 served as the experimental group and XI A1 as the control group, based on existing class assignments and prior academic performance rather than random allocation.

Table 1. Participant characteristics and group comparison design

Characteristic	Experimental group	Control group
Class	XI A2	XI A1
Number of students	26	26
Group allocation	Nonrandom; existing intact class	Nonrandom; existing intact class
Baseline assessment	Pretest	Pretest
Instructional condition	MID, experiments, multimodal instruction	Conventional instruction
Final assessment	Posttest	Posttest
Paired measurements	Pretest and posttest	Pretest and posttest

Note. The student was the unit of measurement, whereas instruction was delivered at the class level. The same pretest and posttest instruments were administered to both groups.

Table 1 summarizes the existing class assignments used in the study. XI A2 was designated the experimental class and XI A1 the control class, with 26 students in each. Measurements were obtained before and after instruction. Both groups completed the same assessment instrument to ensure comparable scoring procedures across conditions.

Students remained in their existing classes without random reassignment. The experimental class received MID instruction integrating experiments and multimodal representations, while the control class received conventional instruction. This arrangement was maintained throughout the study. Baseline scores were subsequently included as a covariate when comparing posttest outcomes.

Implementation of Instructional Treatments

Instruction took place over three meetings totaling nine 40-minute lessons. The topics included the fundamentals and speed of sound waves, vibrating strings

and organ pipes, sound intensity and intensity level, interference, and the Doppler effect. In the experimental class, MID was implemented through lead-in, reconstruction, and production stages incorporating experiments and verbal, visual, auditory, and digital representations. The control class covered the same content for the same duration through conventional instruction.

Table 2. Operational instructional protocol to support replication

Component	Experimental group (multimodal MID)	Control group
Duration	3 meetings; total of nine 40-min lessons	3 meetings; total of nine 40-min lessons
Lead-in	Sound phenomena, guiding questions, links to prior knowledge	Teacher-led introduction and explanation
Reconstruction	Sound experiments, observation, recording, discussion, images, audio/video, worksheets, PhET simulations	Concept explanations, questions and answers, examples, and exercises
Production	Application of concepts, conclusions, communication of observations	Practice questions and discussion of solutions
Example activities	Tin-can telephone, sound observations using water-filled bottles, digital representations	Did not use the same MID experimental sequence
Outcome assessment	Seven-item constructed-response pretest and posttest	Same constructed-response pretest and posttest

Note. The MID stages, sample activities, media types, and comparison-group condition are summarized from the implemented instructional plan.

Table 2 presents the sequence of learning activities used in both groups across three meetings. The experimental group used lead-in to connect sound phenomena with prior knowledge, reconstruction to conduct observation and experiment-based discussions, and production to apply concepts and communicate findings. At each stage, the teacher employed visual and auditory materials and worksheets in line with the instructional objectives.

In the control group, lessons consisted of teacher-led introductions, conceptual explanations, questions and answers, and problem-solving exercises. Both groups studied the same sound-wave topics over nine lessons across three meetings. The measurement procedures were also identical, with a pretest before and a posttest after instruction.

Assessment Instrument and Measurement Procedures

Conceptual understanding was measured using a constructed-response test on sound waves. The initial instrument contained nine items and was piloted with 34 students who had previously studied the topic. Item analysis used Pearson correlations between individual item scores and the total score, and internal consistency was evaluated using Cronbach's alpha. The seven-item test administered at pretest and posttest comprised initial items 1, 2, 3, 6, 7, 8, and 9.

Each response was evaluated using a scoring rubric, and the total score was converted to a 0-100 scale. Cognitive indicators comprised C1 (remembering), C2 (understanding), C3 (applying), C4 (analyzing), C5 (evaluating), and C6 (creating).

Their respective maximum scores were 3, 3, 6, 4, 4, and 4, giving an overall maximum of 24. The same item structure and scoring procedures were used for both assessment occasions.

Figure 1 provides English-language schematic illustrations of the instructional and assessment materials. Panel (a) reconstructs the vibration-and-water observation activity depicted in the experimental worksheet, whereas panel (b) illustrates a contextual question about sound transmission. Both panels are explanatory schematics; the actual study measurements relied on the seven selected test items and the scoring rubric.

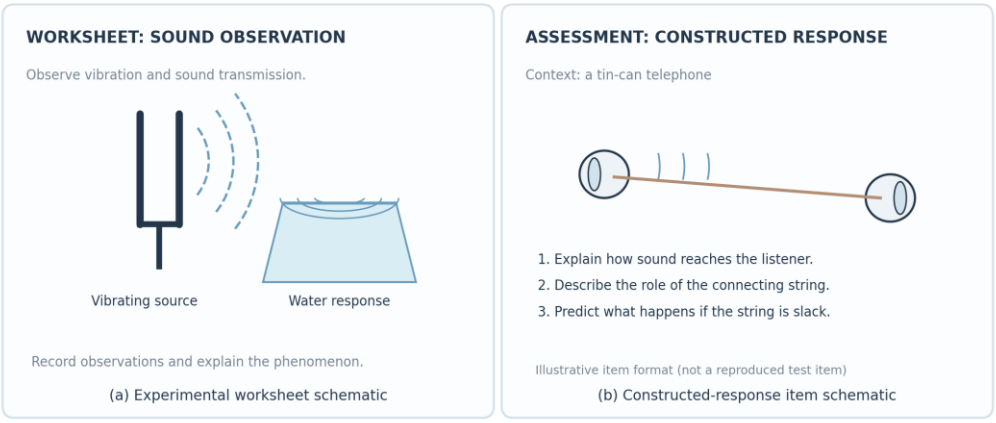


Figure 1. Schematic illustrations of instructional and assessment materials: (a) vibration and water-response activity represented in the experimental worksheet; (b) illustrative constructed-response item on sound transmission.

Panel (a) of Figure 1 schematically represents a vibration and sound observation featured in the experimental worksheet. The activity was used to prompt observation and discussion of the relationships among vibrations, transmission media, and the resulting sound. In the MID reconstruction stage, this type of visual support was accompanied by instructions and records of observations.

Panel (b) of Figure 1 presents a schematic illustration of a contextual constructed-response question using a tin-can telephone. The question demonstrates the connection between an everyday situation and sound transmission through a medium. In the measurement procedure, students answered the seven selected written items; responses were graded using the rubric and converted to scores on a 0-100 scale.

Data Collection Procedures

Data collection comprised preparation, implementation, and data processing. Preparation involved developing instructional materials, worksheets, multimodal resources, test items, and scoring rubrics, followed by instrument piloting. Implementation began with a pretest in both groups, followed by three instructional meetings under the respective conditions, and concluded with a posttest using the same instrument. Responses were scored against the rubric to obtain individual measurements at both time points.

Data Analysis

Descriptive statistics included sample size (n), mean, standard deviation, minimum, and maximum for pretest scores, posttest scores, and individual change

scores (posttest minus pretest). The Shapiro-Wilk test was used to examine score distributions, and Levene's test assessed variance homogeneity across groups. Independent-samples t-tests provided supplementary comparisons of posttest and change scores, with standardized differences reported using Hedges' g. All inferential tests used a significance threshold of .05.

Table 3. Operational definitions and data analysis procedures

Component	Operational definition	Analysis/reporting
Baseline achievement	Pretest score (0-100)	Mean, SD, min-max; covariate
Final achievement	Posttest score (0-100)	Mean, SD; ANCOVA dependent variable
Change score	Posttest - pretest	Mean, SD; independent-samples t-test
Statistical assumptions	Scores, residuals, and group-by-covariate interaction	Shapiro-Wilk, Levene, homogeneity of slopes
Adjusted difference	Group as factor; pretest as covariate	Type II ANCOVA; F, p, partial eta squared, adjusted means, 95% CI
Individual normalized gain	(posttest - pretest)/(100 - pretest)	Group mean normalized gain
Cognitive indicators C1-C6	Percentage of maximum indicator score	Pretest, posttest, and percentage-point difference

Note. Total scores were converted to a 0-100 scale. Normalized gain was calculated for baseline scores below the maximum; changes in C1-C6 indicators were expressed in percentage points.

Table 3 summarizes the analytical treatment of each variable. Pretest and posttest scores were described using means, standard deviations, and ranges. Change scores were calculated by subtracting the pretest from the posttest score for each student. Independent-samples t-tests were used for supplementary between-group comparisons, whereas ANCOVA served as the primary analysis with pretest scores entered as a covariate.

Relative improvement was calculated for each student using normalized gain, defined as the observed score increase divided by the potential increase to the maximum score. Cognitive indicators C1-C6 were analyzed separately by calculating the percentage of the maximum possible score at pretest and posttest. Differences between these percentages were reported in percentage points for each indicator.

In the ANCOVA model, instructional group was specified as a fixed factor, posttest score as the dependent variable, and pretest score as a covariate. Assumptions were assessed through Shapiro-Wilk tests of residual normality, Levene's tests of variance homogeneity, and a group-by-pretest interaction test for homogeneity of regression slopes. Adjusted group means were estimated at a common covariate value and accompanied by 95% confidence intervals.

RESULTS AND DISCUSSION

Validity and Reliability of the Conceptual Understanding Instrument

The pilot instrument comprised nine constructed-response questions administered to 34 students. Pearson product-moment correlations between each item and the total score were used to examine item validity, while Cronbach's alpha

was calculated to assess internal consistency. The final test comprised seven selected items, corresponding to original items 1, 2, 3, 6, 7, 8, and 9.

Table 4. Item validity and reliability of the conceptual understanding instrument

Original item	r	Included in final test
1	0.465	Yes
2	0.443	Yes
3	0.533	Yes
4	0.570	No
5	0.417	No
6	0.511	Yes
7	0.576	Yes
8	0.751	Yes
9	0.672	Yes

Note. Pearson correlations were calculated from 34 responses and the total score for all nine items. The critical r value was .361 (alpha = .05). Cronbach's alpha was .713 for the nine-item test and .683 for the seven selected items. The selection followed the instrument administered in the study.

Item-total Pearson correlations for the nine pilot items ranged from .417 to .751. All coefficients exceeded the critical value of .361 at the .05 significance level for 34 pilot respondents. Table 4 also identifies the seven items included in the final assessment (items 1, 2, 3, 6, 7, 8, and 9) and the two items not retained.

Cronbach's alpha was .713 for the complete nine-item instrument and .683 for the seven selected items. Internal consistency was therefore lower for the final version and slightly below the commonly used .70 guideline. These results warrant caution when interpreting measurement outcomes, particularly indicator-level comparisons based on a limited number of item components.

Descriptive Statistics of Sound-Wave Conceptual Understanding

A total of 52 students completed both the pretest and posttest, comprising 26 students in each group. Descriptive statistics based on individual scores are presented in Table 5.

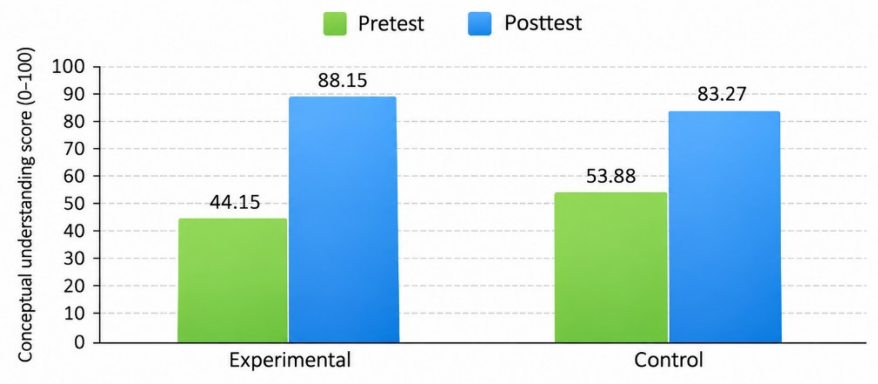
Table 5. Descriptive statistics for pretest and posttest scores

Group	Assessment	n	M	SD	Min.	Max.
Experimental	Pretest	26	44.15	22.52	8	79
	Posttest	26	88.15	8.67	71	100
Control	Pretest	26	53.88	19.27	17	88
	Posttest	26	83.27	10.71	63	100

Note. M = mean; SD = standard deviation. Scores ranged from 0 to 100.

The experimental group obtained a pretest mean of 44.15 (SD = 22.52), whereas the control group obtained 53.88 (SD = 19.27). The control group's mean baseline score was thus 9.73 points higher. This descriptive baseline difference is important when comparing outcomes, because unadjusted posttest comparisons may reflect variation in initial achievement.

Posttest means were 88.15 (SD = 8.67) in the experimental group and 83.27 (SD = 10.71) in the control group. The experimental group scored 4.88 points higher on average and displayed less dispersion in posttest scores. Students in both groups attained the maximum score of 100. The ranges and standard deviations indicate that individual achievement remained variable despite higher group means.



Group	Pretest (*M* ± SD)	Posttest (*M* ± SD)	Increase (Δ)
Experimental	44.15 ± 22.52	88.15 ± 8.67	44.00
Control	53.88 ± 19.27	83.27 ± 10.71	29.38

Figure 2. Comparison of pretest and posttest mean conceptual understanding scores by instructional group.

Figure 2 illustrates increases in both groups. The experimental group's mean rose from 44.15 to 88.15, while the control group's mean increased from 53.88 to 83.27. The ordering of group means reversed between baseline and posttest: the control group began higher, whereas the experimental group finished higher. The figure summarizes the magnitude and direction of descriptive changes without establishing their statistical significance.

The experimental group's mean increased by 44.00 points, compared with 29.38 points in the control group. The descriptive difference in increases was 14.62 points, but this alone does not establish a treatment effect. Inferential comparisons were therefore conducted to distinguish differences in score change from differences in final achievement while accounting for the initial group differences. The figure serves primarily as a concise depiction of change in the two groups.

Changes in Conceptual Understanding Scores

Individual score change was calculated as posttest minus pretest. Descriptive statistics for these changes are reported in Table 6.

Table 6. Descriptive statistics for changes in conceptual understanding scores

Group	n	Mean change	SD	Min.	Max.
Experimental	26	44.00	20.70	8	80
Control	26	29.38	16.58	8	75

Note. Change score = posttest minus pretest; positive values indicate improvement.

The mean change was 44.00 points in the experimental group and 29.38 points in the control group. All paired observations showed positive change, and the experimental group's mean increase exceeded that of the control group by 14.62 points. These descriptive results were calculated from individual change scores rather than an inferential between-group comparison. Statistical tests of the difference in change are reported below.

Individual gains ranged from 8 to 80 points in the experimental group and from 8 to 75 points in the control group. Standard deviations were 20.70 and 16.58, respectively, indicating greater variation in individual gains within the experimental group. The ranges and standard deviations show that improvements were

heterogeneous across students. Group means should therefore be interpreted alongside measures of dispersion rather than as uniform individual responses.

Statistical Assumption Checks

Score distributions were assessed using the Shapiro-Wilk test, and between-group variance homogeneity was examined using Levene's test. The results are reported in Table 7.

Table 7. Normality and variance-homogeneity tests

Assessment	Group	Shapiro-Wilk W	p
Pretest	Experimental	0.923	0.052
	Control	0.945	0.174
Posttest	Experimental	0.910	0.027
	Control	0.939	0.126
Change score	Experimental	0.941	0.145
	Control	0.939	0.129

Note. Levene's test: pretest p = .192; posttest p = .208; change score p = .070. Alpha = .05.

The Shapiro-Wilk tests did not detect significant departures from normality for either group's pretest or change scores. The experimental group's posttest scores departed significantly from normality (p = .027), whereas the control group's posttest scores did not (p = .126). Assumption-check results therefore differed across score types. Subsequent analyses considered the distribution of statistical-model residuals rather than relying solely on the normality of raw scores.

Levene's test yielded p values above .05 for pretest, posttest, and change scores, providing insufficient evidence to reject variance homogeneity. For ANCOVA, residual properties and the homogeneity of regression slopes were examined in addition to raw-score distributions. A nonsignificant Levene test does not establish identical population variances, but indicates that the available data did not detect a statistically significant difference. These results informed the subsequent analyses.

Between-Group Score Comparisons

Independent-samples t-tests were used for supplementary comparisons of posttest and change scores. Two-sided test statistics and standardized effect sizes are presented in Table 8.

Table 8. Independent-samples t-tests comparing the two groups

Variable	Mean diff.	t	df	p	Hedges' g
Posttest	4.88	1.81	50	0.077	0.49
Change score	14.62	2.81	50	0.007	0.77

Note. Differences were calculated as experimental minus control; alpha = .05. Standardized effect sizes are reported as Hedges' g.

The unadjusted posttest difference was not statistically significant, t(50) = 1.81, p = .077. The experimental group's mean posttest score was 4.88 points higher, corresponding to Hedges' g = .49. The standardized effect size describes the mean difference relative to score variability, while the p value quantifies evidence against the null hypothesis. Both measures are reported to avoid evaluating the comparison solely in terms of statistical significance.

The difference in change scores was statistically significant, t(50) = 2.81, p = .007, with Hedges' g = .77. The experimental group gained 14.62 points more, on

average, than the control group. This test addresses differences in score gains rather than final scores after adjustment for initial achievement. The covariance analysis is reported separately to compare posttest achievement while accounting for baseline scores.

Posttest Comparisons Adjusted for Baseline Achievement

An analysis of covariance (ANCOVA) used posttest scores as the dependent variable, instructional group as the factor, and pretest scores as the covariate. The group-by-pretest interaction was not significant, $F(1, 48) = 1.21$, $p = .276$. Model residuals yielded Shapiro-Wilk $W = .965$, $p = .130$, and Levene's test $p = .448$. Results for the main model are presented in Table 9.

Table 9. ANCOVA results for posttest scores

Source	SS	df	MS	F	p	ηp²
Group	587.12	1	587.12	7.58	0.008	0.134
Pretest	947.95	1	947.95	12.23	0.001	0.200
Error	3796.55	49	77.48	–	–	–

Note. SS = sum of squares; MS = mean square; partial eta squared = eta_p^2. Type II sums of squares were used.

ANCOVA identified a significant posttest difference between the experimental and control groups after adjusting for pretest achievement, $F(1, 49) = 7.58$, $p = .008$, partial eta squared (ηp^2) = .134. Instructional group remained significantly associated with final achievement when initial scores were included as a covariate. This result derives from a model incorporating group assignment and pretest achievement simultaneously; the group F statistic therefore assesses the adjusted between-group contrast.

Pretest was also a significant covariate, $F(1, 49) = 12.23$, $p = .001$, partial eta squared = .200. The partial eta-squared value of .134 for the group effect reflects the share of effect-plus-error variance attributable to the group term, not the proportion of total raw posttest variance explained by instruction. Its interpretation is specific to the adjusted model and the residual variance associated with the tested effect.

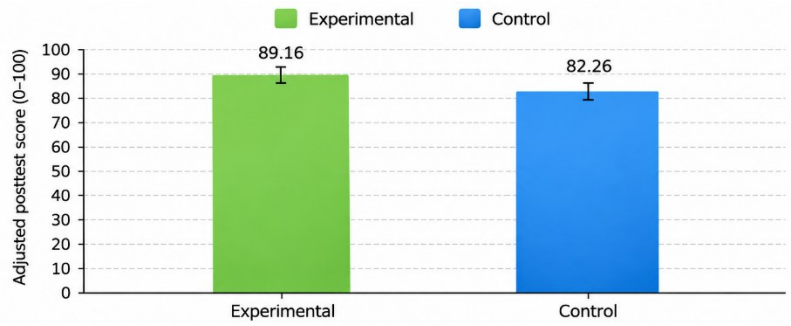
Adjusted posttest means were calculated at the overall mean pretest covariate value to show the magnitude and direction of the between-group contrast. These estimates are given in Table 10.

Table 10. Observed and ANCOVA-adjusted posttest means

Group	n	Observed mean	Adjusted mean
Experimental	26	88.15	89.16
Control	26	83.27	82.26
Difference	–	4.88	6.91

The adjusted posttest mean was 89.16 for the experimental group and 82.26 for the control group. The adjusted mean difference was 6.91 points, exceeding the unadjusted difference of 4.88 points. Both adjusted means were estimated at the same pretest covariate value. The change in the estimated contrast after adjustment indicates that baseline achievement affected the raw comparison. Adjusted means are model-based estimates and should not be confused with observed group means.

The 95% confidence interval for the adjusted mean difference was [1.86, 11.95], excluding zero. This interval is consistent with the statistically significant ANCOVA result. The adjusted means are model estimates rather than directly observed group scores. The confidence interval describes the range of differences compatible with the data and model at the specified confidence level, thereby supplementing the F statistic and p value with information about estimation precision.



Group	Adjusted Mean (*M* ± SE)	95% CI	Difference (Δ)
Experimental	89.16 ± 2.34	[84.46, 93.86]	6.91
Control	82.26 ± 2.29	[77.66, 86.86]	
ANCOVA: *F*(1, 49) = 7.58, *p* = .008; 95% CI of the difference: [1.86, 11.95].			

Figure 3. ANCOVA-adjusted posttest means by instructional group, with 95% confidence intervals.

Figure 3 shows that the adjusted posttest mean was higher in the experimental than in the control group (89.16 vs. 82.26), with both means evaluated at the same pretest covariate value. Error bars indicate the 95% confidence interval around each adjusted mean. The display facilitates a comparison at a common statistical baseline and illustrates the uncertainty associated with each adjusted group mean.

The adjusted mean difference was 6.91 points, with a 95% confidence interval of [1.86, 11.95]. The difference was significant in ANCOVA, F(1, 49) = 7.58, p = .008. The confidence interval for the between-group difference is distinct from the confidence intervals around the two adjusted means shown in the figure. Accordingly, conclusions about between-group differences are based on the estimated contrast and model test, rather than visual overlap of the individual error bars.

Improvement in Conceptual Understanding Based on Normalized Gain

Normalized gain (N-Gain) was calculated from each student's paired pretest and posttest scores and then averaged within groups. Summary values are shown in Table 11.

Table 11. Descriptive statistics for normalized gain in conceptual understanding

Group	n	Mean N-Gain	SD	Category
Experimental	26	0.778	0.183	High
Control	26	0.632	0.206	Moderate

Note. Categories: low, g < .30; moderate, .30 <= g < .70; high, g >= .70.

The experimental group's mean normalized gain was .778 (SD = .183), categorized as high. The control group's mean was .632 (SD = .206), categorized as moderate. Both groups showed positive improvement after instruction. Categories

were assigned using the group mean of individual normalized gains. Because gains were first calculated for each student, these means need not equal normalized gains calculated directly from group mean pretest and posttest scores.

The difference in group mean normalized gains was .146, favoring the experimental group. The experimental group's gain scores also displayed slightly lower variability. This comparison is descriptive and does not establish statistical significance for the difference between normalized gains. Normalized gain summarizes improvement relative to each student's available score range and complements the absolute change-score analysis; these measures differ because normalized gain accounts for the distance from pretest achievement to the maximum possible score.

Changes in Cognitive Indicators C1-C6

Changes across the six cognitive indicators were analyzed using the aggregate indicator scores for 26 students in each group. Scores were divided by the maximum possible score for each indicator and expressed as percentages. Change was calculated as the posttest percentage minus the pretest percentage and reported in percentage points. This procedure avoids undefined individual normalized gains when a student already attained the maximum indicator score at pretest.

Achievement on cognitive indicators C1-C6 was expressed as a percentage of the maximum possible score for each indicator at pretest and posttest. Differences were expressed in percentage points, permitting comparison among indicators with different maximum scores. Group summaries are presented in Table 12.

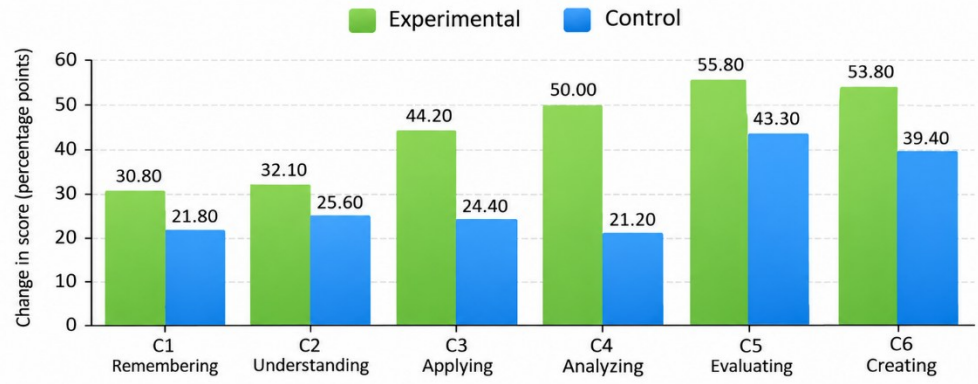
Table 12. Cognitive indicator scores before and after instruction (percentage of maximum possible score)

Indicator	Pre E	Post E	Δ E	Pre C	Post C	Δ C
C1 Remembering	57.7	88.5	30.8	74.4	96.2	21.8
C2 Understanding	59.0	91.0	32.1	65.4	91.0	25.6
C3 Applying	49.4	93.6	44.2	64.1	88.5	24.4
C4 Analyzing	35.6	85.6	50.0	53.8	75.0	21.2
C5 Evaluating	33.7	89.4	55.8	39.4	82.7	43.3
C6 Creating	29.8	83.7	53.8	28.8	68.3	39.4

Note. E = experimental group; C = control group; delta = change in percentage points. Indicator maximum scores: C1 = 3, C2 = 3, C3 = 6, C4 = 4, C5 = 4, and C6 = 4. Values were computed from the indicator-level score matrices for 26 students per group.

Table 12 shows increased percentage scores for all six indicators in both groups. The largest improvement in the experimental group was observed for C5 (evaluating), at 55.8 percentage points, followed by C6 (creating), at 53.8, and C4 (analyzing), at 50.0. The control group's largest improvement was for C5, at 43.3 percentage points. These are descriptive differences and do not imply statistically significant effects for individual indicators.

For C1, the control group improved by 21.8 percentage points, compared with 30.8 points in the experimental group. The experimental group's increase on C4 was 50.0 percentage points versus 21.2 in the control group, producing the largest between-group difference in indicator-level change. Indicator percentages were calculated from aggregate scores and should not be equated with means of individual normalized gains.



Group	Change in score (percentage points)					
	C1	C2	C3	C4	C5	C6
Experimental	30.80	32.10	44.20	50.00	55.80	53.80
Control	21.80	25.60	24.40	21.20	43.30	39.40
Difference (Δ)	9.00	6.50	19.80	28.80	12.50	14.40

Figure 4. Changes in indicator scores (percentage points) across cognitive levels C1-C6.

Figure 4 depicts pretest-to-posttest changes in each indicator for both groups. On C3, the experimental group exhibited greater improvement than the control group, whereas the difference was greatest on C4. Reporting change in percentage points permits comparisons across the six indicators despite their different maximum scores. The figure provides a descriptive view of the magnitude and direction of change.

Percentage-point changes avoid the mathematical problem of undefined individual normalized gains when a maximum score has already been reached at pretest. Figure 4 shows descriptive changes only and does not include inferential testing for each indicator. Thus, statistical interpretation of the instructional comparison remains based on the total-score analyses, particularly ANCOVA.

Discussion

Differences in Sound-Wave Conceptual Understanding after Adjusting for Baseline Achievement

The principal finding was that experiment-based and multimodal Meaningful Instructional Design (MID) was associated with higher sound-wave conceptual understanding than conventional instruction after baseline achievement had been taken into account. ANCOVA indicated a significant between-group difference, $F(1, 49) = 7.58$, $p = .008$, with partial eta squared (ηp^2) = .134. Adjusted posttest means were 89.16 for the experimental group and 82.26 for the control group, yielding a difference of 6.91 points and a 95% confidence interval of [1.86, 11.95]. Thus, the difference in final achievement persisted when the groups were compared at a common pretest covariate value. This finding aligns with the view that learning activities requiring observation, reasoning, and application can support conceptual development. A meta-analysis of active learning (Freeman et al., 2014) found that student engagement in STEM learning was associated with better academic performance than predominantly lecture-based instruction. Although that evidence largely concerns higher education, the underlying role of cognitive engagement provides a useful perspective on sound-wave learning at the secondary level.

The results of the different statistical comparisons also require careful interpretation. The independent-samples t-test for posttest scores was not significant, $t(50) = 1.81$, $p = .077$, whereas the test of change scores was significant, $t(50) = 2.81$, $p = .007$. This contrast should be considered in light of different initial achievement levels. The experimental group's pretest mean was 44.15, below the control group's mean of 53.88; their respective posttest means were 88.15 and 83.27. A comparison of raw posttest scores does not adjust for the baseline difference. ANCOVA instead includes baseline variation as a covariate. Work on statistical causal inference in physics education research (Adlakha & Kuo, 2023) emphasizes the distinction between model-based associations and causal interpretations. Accordingly, the present results support an association between instructional condition and conceptual understanding after baseline adjustment, but do not isolate the separate contributions of MID, experimentation, and multimodal representation.

The group effect size of partial eta squared (ηp^2) = .134 provides information beyond the binary significance decision. It describes a partial association between instructional condition and posttest achievement after accounting for baseline performance. Effect sizes should be considered alongside instructional context, outcome measures, and estimation precision rather than treated as standalone indicators of educational effectiveness. Theobald et al. (2020) showed that the benefits of active learning can be examined through achievement patterns and differences in students' educational experiences. In the present study, the adjusted mean difference, its confidence interval, and the partial effect size jointly provide a more informative picture than the p value alone. The findings are consistent with the educational relevance of instruction that integrates conceptual activity with empirical learning experiences.

Changes in Conceptual Understanding and Relative Improvement

Both groups showed improved sound-wave conceptual understanding, but the magnitudes of change differed. The experimental group gained an average of 44.00 points, compared with 29.38 points in the control group, a between-group difference in change of 14.62 points. This pattern indicates a larger improvement among students experiencing the integrated instructional approach. A plausible interpretation is that these students engaged more extensively in connecting observations, developing explanations, and applying physical concepts. Research on inquiry-based physics laboratory instruction (Thacker, 2023) emphasizes the importance of structuring guided activities so that students use experimental experiences to develop conceptual reasoning rather than simply follow procedures. Alignment between laboratory tasks and conceptual objectives is therefore relevant to interpreting these achievement differences.

Normalized gain offers a complementary description of relative improvement. The experimental group's mean N-Gain was .778 (high), compared with .632 (moderate) in the control group. Because normalized gain measures observed improvement relative to each student's remaining opportunity to improve, it provides different information from absolute score change, particularly when the groups differ at baseline. Research on active engagement and skill development in laboratories suggests that tasks requiring decisions, interpretation, and communication can support learning outcomes (Walsh et al., 2022). The present

pattern is consistent with the possibility that structured experimental experiences helped students strengthen conceptual connections. Nevertheless, normalized gains were used descriptively; the inferential comparison between groups relied on the total-score analyses rather than a separate significance test of N-Gain.

The larger improvement in the experimental group can also be considered in relation to the progression from concrete experiences to conceptual representations. Physics learning requires students to move from noticing a phenomenon to explaining it using accepted principles. Research on analogical scaffolding (Podolefsky & Finkelstein, 2007) indicates that familiar representations can help learners understand abstract physics ideas. Tin-can telephone activities, observations of vibration, and explorations of resonance provide tangible events that can be connected to transmission media, vibrating sources, and wave properties. This connection offers a plausible account of the observed improvements when students are encouraged to compare experimental evidence with theoretical explanations. The proposed mechanism, however, is based on instructional characteristics and related literature rather than direct measurement of cognitive processes during each activity.

The Role of Meaningful Instructional Design in Organizing Conceptual Learning

The lead-in, reconstruction, and production stages of MID provide a structure for linking prior knowledge to new learning experiences. During lead-in, everyday sound phenomena activate students' initial ideas about sound sources, media, and propagation. Reconstruction guides students to observe physical events and relate evidence to conceptual explanations. Production provides opportunities to apply and communicate understanding through problem solving and interpretation of observations. This sequence is compatible with the conceptual-resources perspective, in which students' existing ideas are treated as potential resources for scientific reasoning. Robertson et al. (2023) emphasize identifying productive elements in students' responses rather than merely classifying answers as correct or incorrect. This perspective provides a rationale for activating prior knowledge within MID.

Students' intuitive understanding of sound is often formed through experiences such as hearing voices, observing vibrating objects, and using simple communication devices. Such experiences can produce explanations that are not fully scientific but still contain potentially useful ideas. Eshach and Schwartz (2006) showed that students can interpret sound through materialistic conceptions inconsistent with the scientific explanation of propagation. Reiner et al. (2000) similarly documented substance-based reasoning in everyday physics. In this context, lead-in can surface initial ideas for further examination, while reconstruction creates opportunities to relate those ideas to observed evidence. This organization serves the conceptual aim of strengthening connections between intuitive experiences and scientific models.

Developing initial ideas into scientific explanations also depends on how students draw on existing knowledge. In research on electric circuits, Bauman et al. (2024) found that student responses could contain productive conceptual resources even when final answers were not fully consistent with accepted models. Although the domain differs from sound waves, the study supports using elements of initial

reasoning as a starting point for conceptual development. The observation and discussion activities during reconstruction may have provided similar opportunities in the present study. By observing source vibrations, medium behavior, and changing experimental conditions, students were encouraged to connect concrete phenomena to physical mechanisms. This interpretation may help explain the different improvement patterns compared with instruction centered on explanations and exercises.

Production is conceptually relevant because it requires students to use knowledge after observation and explanation building. Through problem solving, presentation, and summarizing, learners organize information into defensible explanations. Research on rubrics for assessing thinking skills in free-response questions (Al-Salmani & Thacker, 2021) indicates that question demands and response formats can reveal different levels of cognitive processing. Subsequent work on assessing thinking through free-response examination tasks likewise considers how written answers reveal reasoning quality rather than only final-answer correctness (Al-Salmani et al., 2023). From this perspective, production activities are relevant to explaining conceptual relationships in sound-wave problems. The observed achievement gains can be interpreted in relation to the integrated MID sequence without assuming that each stage contributed a separately measured effect.

Integrating Experiments and Multimodal Representations in Sound-Wave Instruction

Multimodal representations are especially relevant to sound waves because many features of these phenomena cannot be observed directly. Students can hear sound and observe vibrating sources, but pressure variations, propagation patterns, and quantitative relationships require models and other representations. Images, simulations, graphs, spoken explanations, and auditory experiences offer different channels through which the same phenomenon can be understood. Svensson and Campos (2022) examined the communicative and meaning-making functions of representations in physics education. In this study, multimodal representations may have helped students relate observed phenomena to physical concepts. Representations therefore served not only to display phenomena but also to structure information and coordinate scientific explanations.

How representations are presented matters as much as their presence. Susac et al. (2023) found that verbal, pictorial, and graphical formats can elicit different patterns of information processing in conceptual questions, even when achievement differences are not always significant. Thus, adding images does not necessarily improve understanding unless students can interpret them. Lichtenberger et al. (2024) likewise reported that a gradual shift from concrete to abstract representations does not consistently outperform simultaneous presentation. These findings suggest that multimodal learning depends on the coordination of representations and their alignment with conceptual objectives. Accordingly, the observed pattern in the experimental group is better interpreted in relation to integrated instructional activities than the number of media employed.

The relationship between visual representations and problem solving is supported by work in other physics domains. Vignal and Wilcox (2022) showed that diagrams created by students may serve different purposes in problem solving and

in communicating explanations. Representations are therefore part of reasoning rather than merely final products. Marshman et al. (2024), studying quantum states, also illustrated how coordinating multiple representations can support understanding of abstract material. Although those studies did not examine sound waves directly, they highlight the value of asking students to interpret, compare, and translate information across representational formats. Similar demands arise when MID learners connect experimental records, images, simulations, and written explanations.

Physical experiments and digital simulations can provide complementary experiences. Physical experiments enable direct interaction with phenomena, while simulations allow learners to visualize and manipulate conditions that may be difficult to observe. Wörner et al. (2022) concluded that combinations of real and virtual experiments can support learning, with effectiveness depending on activity design and instructional organization. In the present study, simple experiments and PhET simulations combined physical experiences with digital representations. Diwakar et al. (2023) found that virtual laboratory use is also related to motivational and interactional dimensions of learning. This provides relevant context for digital learning environments, although motivation cannot be inferred to mediate the present score differences because it was not directly measured.

Representational quality and visual demands are also relevant to conceptual interpretation. Wang and Wei (2024) reported that visual physics questions can involve different levels of cognitive processing. In sound-wave contexts, illustrations of resonance, wave patterns, or changes in source conditions may require identification, interpretation, and analysis. Rangkuti and Karam (2022) showed that students may struggle to interpret graphical representations of wave-pulse propagation. These findings reinforce that visual representations are not necessarily cognitively simple. The present improvement pattern may be associated with activities that encouraged students to use representations to construct explanations, rather than only to view pictures or simulations.

Variation in Changes across Cognitive Indicators C1-C6

Indicator-level analyses showed improvement across all six cognitive processes in both groups, with different magnitudes across indicators and conditions. In the experimental group, the largest increase occurred for C5 (evaluating), at 55.8 percentage points, followed by C6 (creating), at 53.8, and C4 (analyzing), at 50.0. The control group's largest increase was also for C5, at 43.3 percentage points. The largest between-group difference in change occurred for C4, at 28.8 percentage points. Thus, gains in conceptual understanding were not uniform across indicators. Research on thinking-skill assessment rubrics (Al-Salmani & Thacker, 2021) emphasizes attending to the different cognitive demands of students' responses. This perspective is useful because a gain in total score may reflect changes in processes ranging from recall to explanation and evaluation.

The larger changes for C4, C5, and C6 in the experimental group may be interpreted in relation to activities involving observation analysis, conclusion drawing, and conceptual application. During reconstruction, students observed phenomena and discussed relationships between evidence and physical principles. During production, they applied knowledge to problems and communicated explanations. Walsh et al. (2022) showed that laboratory activities involving decision

making and communicating results are associated with critical-thinking skills and views about experimentation. Although their measures and setting differed, these instructional principles help explain why changes may extend beyond remembering and understanding. The present indicator differences remain descriptive and should not be interpreted as statistically established effects at each cognitive level.

On C1 and C2, both groups obtained relatively high posttest percentages, suggesting that foundational knowledge and direct conceptual understanding developed under both instructional conditions. More pronounced variation appeared on higher-level indicators, particularly when students had to integrate several kinds of conceptual information within a problem. Volfson et al. (2021) suggest that diagnostic response patterns may reflect more complex mental-model configurations than simple correct/incorrect classifications. Accordingly, changes across C1-C6 complement total scores by describing different cognitive demands, but do not identify individual transitions between mental models. These results reinforce the value of interpreting total achievement alongside the conceptual demands represented in the assessment.

Differences across cognitive indicators may also reflect the interplay between representations and problem contexts. Goodhew et al. (2019) showed that students draw on diverse conceptual resources when reasoning about mechanical-wave propagation. The multiple-representations framework developed by Ainsworth (2006) likewise emphasizes that different representations can complement one another, constrain interpretations, or support deeper understanding. In sound-wave tasks, analyzing and evaluating may require integrating information about the source, transmission medium, and propagation properties. The observed C1-C6 patterns can therefore be understood as reflecting differences in conceptual demands. Larger gains on higher-level indicators do not establish equal mastery of all related reasoning skills, but describe changes within the measured cognitive categories.

Synthesis of Findings and Contributions to Physics Education

The study provides evidence that experiment-based and multimodal MID is associated with stronger sound-wave conceptual understanding after adjustment for initial achievement. This pattern appears in adjusted mean differences, individual score gains, and cognitive-indicator profiles. These measures address complementary questions: ANCOVA compares final scores after baseline adjustment; absolute changes and normalized gains describe improvement; and C1-C6 analyses characterize differences across cognitive demands. Research on active learning (Freeman et al., 2014) and inquiry-oriented laboratory instruction (Thacker, 2023) supports the relevance of engaging students directly in physics learning. The present contribution is to examine these principles within an MID framework integrating activation of prior knowledge, experimentation, multimodal representation, and conceptual application in sound-wave instruction.

The findings further support treating multimodal instruction as an organized learning system rather than an accumulation of separate media. Simulations, diagrams, verbal explanations, and experiments serve pedagogical purposes when each representation explains a particular aspect of a phenomenon and is connected through student activity. Multimedia learning research (Mayer & Moreno, 2002)

emphasizes aligning verbal and visual support with information-processing demands. Relationships across representations can also shape how learners interpret physics phenomena, as illustrated by Susac et al. (2023) and Lichtenberger et al. (2024). The pedagogical contribution of this study therefore lies in organizing experiences that connect observable sound phenomena, the scientific representations required to explain them, and opportunities to apply conceptual knowledge.

The study also supports viewing sound-wave learning as the development of relationships among concepts, not simply an increase in correct answers. Difficulties in understanding sound involve the interaction of intuitive experiences, conceptual categories, and scientific models. Eshach and Schwartz (2006) documented the influence of materialistic conceptions on sound reasoning, whereas Robertson et al. (2023) emphasized using students' initial ideas as productive resources for scientific understanding. The present findings indicate the relevance of learning experiences that connect concrete situations, observations, and scientific representations within a coherent sequence. This contributes empirical evidence about integrated approaches to sound-wave instruction in secondary education, particularly when achievement is examined through both adjusted group comparisons and changes across cognitive-process indicators.

CONCLUSION

This study found that experiment-based and multimodal Meaningful Instructional Design (MID) was associated with higher sound-wave conceptual understanding than conventional instruction after adjustment for initial achievement. ANCOVA indicated a significant between-group difference, $F(1, 49) = 7.58$, $p = .008$, with partial eta squared (ηp^2) = .134. The adjusted posttest mean was 89.16 in the experimental group and 82.26 in the control group, a difference of 6.91 points with a 95% confidence interval of [1.86, 11.95]. These results support the relevance of MID integrating experiments and multimodal representations in sound-wave instruction when baseline achievement is considered.

Both groups improved, although the magnitude of improvement differed. The experimental group gained an average of 44.00 points, compared with 29.38 points in the control group. The between-group difference in change scores was significant, $t(50) = 2.81$, $p = .007$, with Hedges' $g = .77$. Mean normalized gain was also higher in the experimental group (.778, high) than in the control group (.632, moderate). All C1-C6 cognitive indicators improved. The experimental group's largest changes were in C5 (evaluating; 55.8 percentage points), C6 (creating; 53.8), and C4 (analyzing; 50.0). The largest between-group difference in indicator change occurred for C4 (28.8 percentage points). These patterns suggest that conceptual gains extended across cognitive processes but were uneven in magnitude.

Substantively, the study contributes evidence on sound-wave instruction that integrates prior-knowledge activation, experimentation, multimodal representation, and conceptual application within the MID framework. Its lead-in, reconstruction, and production stages provide an instructional sequence linking concrete experiences with conceptual explanations and problem solving. MID, experiments, and multimodal representations should be viewed as components of an integrated approach rather than independent elements whose separate effects

were estimated. Pedagogically, sound-wave instruction may benefit from explicit connections among empirical phenomena, scientific representations, and students' reasoning activities across cognitive levels.

LIMITATIONS

This study involved only two intact classes, with one class assigned to each instructional condition. Consequently, differences associated with the instructional condition cannot be fully disentangled from class-level characteristics, and adjusting for pretest scores does not eliminate all potential confounding. The intervention lasted three meetings, so the persistence of the observed gains was not evaluated. The final seven-item assessment had a Cronbach's alpha of .683, and indicator-level outcomes were based on relatively limited item coverage. These considerations qualify the generalizability of the findings and the precision of conclusions about individual cognitive indicators. The study assessed MID, experiments, and multimodal representations as a combined instructional condition rather than estimating their separate contributions.

RECOMMENDATIONS

Based on these findings, experiment-based and multimodal MID may be considered as an instructional option for supporting sound-wave conceptual understanding in secondary school physics. Teachers are encouraged to organize lead-in, reconstruction, and production stages systematically by activating prior knowledge, conducting physical experiments, using digital simulations, and coordinating verbal, visual, and auditory representations. Learning activities should help students relate empirical phenomena to physical principles, particularly in analyzing, evaluating, and creating. Future studies should examine the approach across other physics topics and student populations, use repeated measurements to investigate the persistence of conceptual understanding, and compare alternative ways of integrating experiments and multimodal representations within MID. Further development of instruments that probe conceptual reasoning in greater depth would also provide more comprehensive evidence about changes in students' understanding

ACKNOWLEDGEMENT

The authors would like to express their sincere appreciation to the school administration, physics teachers, and students who participated in this study. Their cooperation and participation supported the implementation of the instructional activities and the collection of research data. The authors also acknowledge the academic support provided by the Physics Education Department, Universitas Mataram, Indonesia.

REFERENCES

Adlakha, V., & Kuo, E. (2023). Critical issues in statistical causal inference for observational physics education research. *Physical Review Physics Education Research*, *19*, 020160. https://doi.org/10.1103/PhysRevPhysEducRes.19.020160

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

Al-Salmani, F., Johnson, J., & Thacker, B. (2023). Assessing thinking skills in free-response exam problems: Pandemic online and in-person. *Physical Review Physics Education Research*, *19*, 010131. https://doi.org/10.1103/PhysRevPhysEducRes.19.010131

Al-Salmani, F., & Thacker, B. (2021). Rubric for assessing thinking skills in free-response exam problems. *Physical Review Physics Education Research*, *17*, 010135. https://doi.org/10.1103/PhysRevPhysEducRes.17.010135

Banda, H. J., & Nzabahimana, J. (2021). Effect of integrating physics education technology simulations on students' conceptual understanding in physics: A review of literature. *Physical Review Physics Education Research*, *17*, 023108. https://doi.org/10.1103/PhysRevPhysEducRes.17.023108

Barniol, P., & Zavala, G. (2016). Mechanical waves conceptual survey: Its modification and conversion to a standard multiple-choice test. *Physical Review Physics Education Research*, *12*, 010107. https://doi.org/10.1103/PhysRevPhysEducRes.12.010107

Bauman, L. C., Hansen, B., Goodhew, L. M., & Robertson, A. D. (2024). Student conceptual resources for understanding electric circuits. *Physical Review Physics Education Research*, *20*, 020128. https://doi.org/10.1103/PhysRevPhysEducRes.20.020128

Chi, M. T. H., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, *36*(1), 1–61. https://doi.org/10.1111/j.1551-6709.2011.01207.x

Diwakar, S., Kolil, V. K., Francis, S. P., & Achuthan, K. (2023). Intrinsic and extrinsic motivation among students for laboratory courses: Assessing the impact of virtual laboratories. *Computers & Education*, *198*, 104758. https://doi.org/10.1016/j.compedu.2023.104758

Eshach, H. (2014). Development of a student-centered instrument to assess middle school students' conceptual understanding of sound. *Physical Review Special Topics - Physics Education Research*, *10*, 010102. https://doi.org/10.1103/PhysRevSTPER.10.010102

Eshach, H., Lin, T.-C., & Tsai, C.-C. (2016). Taiwanese middle school students' materialistic concepts of sound. *Physical Review Physics Education Research*, *12*, 010119. https://doi.org/10.1103/PhysRevPhysEducRes.12.010119

Eshach, H., Lin, T.-C., & Tsai, C.-C. (2018). Misconception of sound and conceptual change: A cross-sectional study on students' materialistic thinking of sound. *Journal of Research in Science Teaching*, *55*(5), 664–684. https://doi.org/10.1002/tea.21435

Eshach, H., & Schwartz, J. L. (2006). Sound stuff? Naive materialism in middle-school students' conceptions of sound. *International Journal of Science Education*, *28*(7), 733–764. https://doi.org/10.1080/09500690500277938

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis.

Review of Educational Research, *82*(3), 300–329. https://doi.org/10.3102/0034654312457206

Goodhew, L. M., Robertson, A. D., Heron, P. R. L., & Scherr, R. E. (2019). Student conceptual resources for understanding mechanical wave propagation. *Physical Review Physics Education Research*, *15*, 020127. https://doi.org/10.1103/PhysRevPhysEducRes.15.020127

Goodhew, L. M., Robertson, A. D., Heron, P. R. L., & Scherr, R. E. (2021). Students' context-sensitive use of conceptual resources: A pattern across different styles of question about mechanical waves. *Physical Review Physics Education Research*, *17*, 010137. https://doi.org/10.1103/PhysRevPhysEducRes.17.010137

Hrepic, Z., Zollman, D. A., & Rebello, N. S. (2010). Identifying students' mental models of sound propagation: The role of conceptual blending in understanding conceptual change. *Physical Review Special Topics - Physics Education Research*, *6*, 020114. https://doi.org/10.1103/PhysRevSTPER.6.020114

Kohl, P. B., & Finkelstein, N. D. (2005). Student representational competence and self-assessment when solving physics problems. *Physical Review Special Topics - Physics Education Research*, *1*, 010104. https://doi.org/10.1103/PhysRevSTPER.1.010104

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, *86*(3), 681–718. https://doi.org/10.3102/0034654315627366

Lichtenberger, A., Kokkonen, T., & Schalk, L. (2024). Learning with multiple external representations in physics: Concreteness fading versus simultaneous presentation. *Journal of Research in Science Teaching*, *61*(9), 2258–2290. https://doi.org/10.1002/tea.21947

Marshman, E., Maries, A., & Singh, C. (2024). Using multiple representations to improve student understanding of quantum states. *Physical Review Physics Education Research*, *20*, 020152. https://doi.org/10.1103/PhysRevPhysEducRes.20.020152

Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, *13*(2), 125–139. https://doi.org/10.1016/S0959-4752(02)00016-6

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*(1), 107–119. https://doi.org/10.1016/S0959-4752(01)00018-4

Podolefsky, N. S., & Finkelstein, N. D. (2006). Use of analogy in learning physics: The role of representations. *Physical Review Special Topics - Physics Education Research*, *2*, 020101. https://doi.org/10.1103/PhysRevSTPER.2.020101

Podolefsky, N. S., & Finkelstein, N. D. (2007). Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies. *Physical Review Special Topics - Physics Education Research*, *3*, 020104. https://doi.org/10.1103/PhysRevSTPER.3.020104

Rangkuti, M. A., & Karam, R. (2022). Conceptual challenges with the graphical representation of the propagation of a pulse in a string. *Physical Review*

Physics Education Research, *18*, 020119. https://doi.org/10.1103/PhysRevPhysEducRes.18.020119

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naïve physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, *18*(1), 1–35. https://doi.org/10.1207/S1532690XCI1801_01

Robertson, A. D., Goodhew, L. M., Bauman, L. C., Hansen, B., & Alesandrini, A. T. (2023). Identifying student conceptual resources for understanding physics: A practical guide for researchers. *Physical Review Physics Education Research*, *19*, 020138. https://doi.org/10.1103/PhysRevPhysEducRes.19.020138

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*(2), 227–237. https://doi.org/10.1016/S0959-4752(02)00022-1

Susac, A., Planinic, M., Bubic, A., Jelicic, K., & Palmovic, M. (2023). Effect of representation format on conceptual question performance and eye-tracking measures. *Physical Review Physics Education Research*, *19*, 020114. https://doi.org/10.1103/PhysRevPhysEducRes.19.020114

Svensson, K., & Campos, E. (2022). Comparison of two semiotic perspectives: How do students use representations in physics?. *Physical Review Physics Education Research*, *18*, 020120. https://doi.org/10.1103/PhysRevPhysEducRes.18.020120

Thacker, B. (2023). Inquiry-based experimental physics: Twenty years of an evidence-based, laboratory-based physics course for algebra-based physics students. *Physical Review Physics Education Research*, *19*, 020116. https://doi.org/10.1103/PhysRevPhysEducRes.19.020116

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., . . . Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483. https://doi.org/10.1073/pnas.1916903117

Tongchai, A., Sharma, M. D., Johnston, I. D., Arayathanitkul, K., & Soankwan, C. (2011). Consistency of students' conceptions of wave propagation: Findings from a conceptual survey in mechanical waves. *Physical Review Special Topics - Physics Education Research*, *7*, 020101. https://doi.org/10.1103/PhysRevSTPER.7.020101

Vignal, M., & Wilcox, B. R. (2022). Investigating unprompted and prompted diagrams generated by physics majors during problem solving. *Physical Review Physics Education Research*, *18*, 010104. https://doi.org/10.1103/PhysRevPhysEducRes.18.010104

Volfson, A., Eshach, H., & Ben-Abu, Y. (2018). Development of a diagnostic tool aimed at pinpointing undergraduate students' knowledge about sound and its implementation in simple acoustic apparatuses' analysis. *Physical Review Physics Education Research*, *14*, 020127. https://doi.org/10.1103/PhysRevPhysEducRes.14.020127

Volfson, A., Eshach, H., & Ben-Abu, Y. (2019). Introducing the idea of entropy to the ontological category shift theory for conceptual change: The case of heat and

sound. *Physical Review Physics Education Research*, *15*, 010143. https://doi.org/10.1103/PhysRevPhysEducRes.15.010143

Volfson, A., Eshach, H., & Ben-Abu, Y. (2021). Preliminary development of a simple statistical tool for estimating mental model states from a diagnostic test. *Physical Review Physics Education Research*, *17*, 023105. https://doi.org/10.1103/PhysRevPhysEducRes.17.023105

Volfson, A., Eshach, H., & Ben-Abu, Y. (2022). History of science based dialogues on sound waves: From sound atoms to phonons. *Physical Review Physics Education Research*, *18*, 010123. https://doi.org/10.1103/PhysRevPhysEducRes.18.010123

Walsh, C., Lewandowski, H. J., & Holmes, N. G. (2022). Skills-focused lab instruction improves critical thinking skills and experimentation views for all students. *Physical Review Physics Education Research*, *18*, 010128. https://doi.org/10.1103/PhysRevPhysEducRes.18.010128

Wang, C., & Wei, B. (2024). Analysis of visual-based physics questions of the senior high school entrance examination in China. *Physical Review Physics Education Research*, *20*, 010112. https://doi.org/10.1103/PhysRevPhysEducRes.20.010112

Wang, X.-M., Yu, X.-H., Yu, D.-D., Hwang, G.-J., & Lan, M. (2025). Does combining real and virtual experiments improve learning achievement in physics? Evidence from a meta-analysis (2001–2021). *Educational Research Review*, *46*, 100661. https://doi.org/10.1016/j.edurev.2024.100661

Wittmann, M. C., Steinberg, R. N., & Redish, E. F. (2003). Understanding and affecting student reasoning about sound waves. *International Journal of Science Education*, *25*(8), 991–1013. https://doi.org/10.1080/09500690305024

Wörner, S., Kuhn, J., & Scheiter, K. (2022). The best of two worlds: A systematic review on combining real and virtual experiments in science education. *Review of Educational Research*, *92*(6), 911–952. https://doi.org/10.3102/00346543221079417

Xie, L., Liu, Q., Lu, H., Wang, Q., Han, J., Feng, X., & Bao, L. (2021). Student knowledge integration in learning mechanical wave propagation. *Physical Review Physics Education Research*, *17*, 020122. https://doi.org/10.1103/PhysRevPhysEducRes.17.020122

Zeng, L., Smith, C., Poelzer, G. H., Rodriguez, J., Corpuz, E., & Yanev, G. (2014). Illustrations and supporting texts for sound standing waves of air columns in pipes in introductory physics textbooks. *Physical Review Special Topics - Physics Education Research*, *10*, 020110. https://doi.org/10.1103/PhysRevSTPER.10.020110