Journal of Science and Mathematics Education
Vol. 2 No. 2, June 2026, pp. 58-65
E-ISSN 3090-0336

JOSME Journal of Science and Mathematics Education

AI-Powered Personalized Learning in Indonesian Senior High School Chemistry Education: Evidence from a Quasi-Experimental Study

Reza Firmansyah[1*], Novia Kurniawati[2], Bagas Dwi Prasetyo[3]
[123]Department of Chemistry Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Indonesia.
* Corresponding Author: reza.firmansyah@unesa.ac.id

ARTICLE INFO

Article history
……………………………….
Received : March 21, 2025
Revised : April 24, 2026
Accepted : June 22, 2026
Published: June 23, 2026

Keywords
AI-Powered Learning
Ai-Powered Adaptive Learning Platform
Chemistry Education
Academic Achievement
Personalized Learning

ABSTRACT

This study investigated the effectiveness of an AI-powered AI-powered adaptive learning platform in improving students' academic achievement in chemistry and explored their overall learning experience. The research was conducted at SMA Negeri 3 Surabaya, involving 72 students of class XI MIPA in the 2023/2024 academic year. A quasi-experimental pretest-posttest control group design was employed. The experimental group (XI MIPA 2, n = 36) used an AI-powered adaptive learning platform incorporating artificial intelligence algorithms to personalize learning pathways, whereas the control group (XI MIPA 1, n = 36) received conventional instruction. Academic achievement was measured using a validated chemistry knowledge test covering reaction rates, chemical equilibrium, and thermochemistry, while learning experience was assessed through a structured questionnaire. The results showed that the experimental group achieved significantly higher posttest scores (mean = 83.7) than the control group (mean = 72.4), with a statistically significant difference ($t = 6.43$, $p < 0.001$) and a large effect size ($d = 1.02$). The experimental group also obtained a higher N-Gain score (0.61, moderate) than the control group (0.35, low). In addition, students in the experimental group reported highly positive learning experiences across all measured dimensions. These findings suggest that AI-powered personalized learning has strong potential to enhance chemistry achievement and student engagement in Indonesian senior high school settings.

How to cite: Firmansyah, R., Kurniawati, N., & Prasetyo, B., D. (2026). AI-Powered Personalized Learning in Indonesian Senior High School Chemistry Education: Evidence from a Quasi-Experimental Study. *Journal of Science and Mathematics Education*. *2*(2). 58-65. https://doi.org/10.70716/josme.v2i2.595

INTRODUCTION

The integration of artificial intelligence (AI) into education has emerged as one of the most transformative developments in the 21st century, reshaping how learning is designed, delivered, and experienced. Unlike traditional one-size-fits-all instructional models, AI-powered systems are capable of analyzing individual learner data in real time, identifying knowledge gaps, adjusting the difficulty and sequence of content, and providing targeted feedback tailored to each student's unique needs (Holmes et al., 2019). This capacity for intelligent personalization aligns with long-standing theoretical aspirations in educational psychology, particularly Bloom's (1984) seminal finding that one-on-one tutoring produces learning gains two standard deviations above conventional classroom instruction—a benchmark he termed the "2 sigma problem." AI-driven AI-powered AI-powered adaptive learning platforms represent a scalable technological attempt to approximate this level of individualized instruction for entire classrooms.

In the Indonesian educational context, the adoption of educational technology has accelerated significantly following the disruptions caused by the COVID-19 pandemic (Selvaraj et al., 2021). Schools across the country were compelled to transition rapidly to online and hybrid learning environments, exposing both the potential and the limitations of existing digital infrastructure. The national Kurikulum Merdeka (Kemdikbud, 2022) further emphasizes student-centered learning, differentiated instruction, and the development of higher-order thinking skills—all goals that are theoretically well-served by adaptive AI systems. Despite this enabling policy environment, empirical research on the practical implementation and effectiveness of AI-powered AI-powered AI-powered adaptive learning platforms in Indonesian secondary chemistry education remains sparse, creating a significant knowledge gap that the present study seeks to address.

Chemistry is widely recognized as one of the most conceptually challenging subjects in secondary education. Its abstract nature—involving atomic-level phenomena, mathematical relationships, and complex mechanistic processes—frequently leads to significant disparities in student understanding within the same classroom (Chou & Shih, 2020). These disparities are particularly evident in topics such as reaction rates, chemical equilibrium, and thermochemistry, where students must integrate conceptual understanding with mathematical problem-solving skills. In conventional instruction, teachers often face difficulties in simultaneously addressing the needs of students who are struggling with foundational concepts while also challenging those who have already mastered them. An AI-powered AI-powered AI-powered adaptive learning platform, by contrast, can dynamically route individual students through differentiated content pathways based on their demonstrated performance, thereby continuously optimizing the instructional experience for each learner (Aleven et al., 2017).

Personalized learning, broadly defined as the tailoring of instruction to individual student needs, preferences, and learning trajectories, has received growing empirical support as a strategy for improving academic outcomes. Pane et al. (2017) conducted a large-scale evaluation of personalized learning schools in the United States and found that students in such environments made significantly greater gains in mathematics and reading compared to matched comparison schools. At the classroom level, Walkington (2013) demonstrated that personalizing algebra problems to align with students' personal interests led to significantly improved performance on transfer tasks. Within the domain of AI-supported instruction specifically, Ma et al. (2014) conducted a meta-analysis of 107 studies on intelligent tutoring systems and found a mean effect size of $d = 0.76$ compared to conventional classroom instruction—a substantial and practically meaningful advantage.

AI-powered adaptive learning platforms differ from static e-learning resources in that they incorporate algorithmic mechanisms to continuously monitor learner behavior, assess knowledge states, and modify the instructional pathway accordingly. These mechanisms typically include item response theory (IRT) models for ability estimation, Bayesian knowledge tracing for tracking concept mastery over time, and recommendation algorithms for selecting optimal next steps in the learning sequence (Baker & Inventado, 2014). Leading platforms such as Khan Academy, Carnegie Learning's MATHia, and DreamBox have demonstrated efficacy in various school subjects, but their application to chemistry education in Indonesian senior high schools has not been systematically documented. Given the distinct curricular structure, pedagogical culture, and learner characteristics of the Indonesian educational system, direct extrapolation from international findings must be approached with caution.

The learning experience dimension is equally important in evaluating AI-powered educational tools. Beyond measurable cognitive outcomes, the extent to which students find AI-powered AI-powered adaptive learning platforms engaging, motivating, and user-friendly determines whether such tools will be sustained in practice or abandoned after initial novelty fades (Roll & Wylie, 2016). Research by Zawacki-Richter et al. (2019), based on a systematic review of 146 studies on AI in higher education, found that student engagement and perceived usefulness consistently mediated the relationship between AI tool features and learning outcomes. Hwang and Tu (2021) similarly emphasized the importance of evaluating both affective and cognitive dimensions of AI-mediated learning experiences. In the Indonesian secondary school context, where student motivation in science subjects is a recognized challenge (OECD, 2023), understanding how AI-driven personalization influences learner experience is of practical significance.

Although previous studies have shown that AI-supported and adaptive learning technologies can improve student achievement, most of this evidence has been generated in mathematics, language learning, and higher education contexts. Research focusing specifically on the use of AI-powered AI-powered adaptive learning platforms in secondary school chemistry remains limited, particularly in developing-country settings such as Indonesia. In addition, prior studies have predominantly emphasized cognitive outcomes, while giving less attention to students' learning experiences, such as engagement, motivation, perceived usefulness, and ease of use. Therefore, there is still limited empirical evidence on how AI-powered personalized learning simultaneously affects both academic achievement and learning experience in Indonesian senior high school chemistry education. This study addresses that gap by examining the effectiveness of an AI-driven AI-powered adaptive learning platform in improving chemistry achievement while also evaluating students' learning experiences in an Indonesian secondary school context.

The novelty of this study is explicitly reflected in three main contributions. First, this research extends the literature on AI-adaptive learning by examining its implementation specifically in Indonesian senior high school chemistry education, a context that remains underrepresented in prior studies, which have largely focused on mathematics, language learning, or higher education. Second, unlike many previous studies that primarily emphasize cognitive performance, this study provides a more comprehensive evaluation by investigating both academic achievement and students' learning experience. Third, this study contributes empirical evidence from an Indonesian public school context, thereby enriching current understanding of how AI-powered AI-powered adaptive learning platforms function in emerging educational systems with distinct curricular, technological, and pedagogical characteristics. In this way, the present study does not merely replicate previous findings, but contextualizes and broadens them within secondary chemistry education in Indonesia.

Building on this theoretical and empirical foundation, the present study was conducted at SMA Negeri 3 Surabaya, a public senior high school in the city of Surabaya, East Java. The school was selected as the research site due to its adequate ICT infrastructure, its institutional interest in innovative pedagogical approaches, and the representativeness of its student population in terms of academic heterogeneity. The study aimed to (1) examine the effect of an AI-powered AI-powered adaptive learning platform on students' academic achievement in class XI MIPA, and (2) explore the quality of students' learning experience when engaging with the platform. The findings are expected to contribute both to the theoretical understanding of AI-mediated personalization in science education and to the practical knowledge base available to chemistry teachers and educational technology developers in Indonesia and comparable emerging-economy contexts.

METHOD

Research Design

This study adopted a quasi-experimental design with a pretest-posttest control group structure. The quasi-experimental approach was selected because random assignment of students to groups was not feasible given the intact class structures at the research site (Fraenkel et al., 2012). The independent variable was the type of instruction (AI-powered adaptive learning versus conventional instruction), and the dependent variables were students' academic achievement in chemistry and their reported learning experience. Pretest equivalence between groups was established prior to the intervention through statistical comparison of baseline scores.

Research Location and Participants

The study was conducted at SMA Negeri 3 Surabaya, located at Jalan Praban No. 4, Genteng, Surabaya, East Java, Indonesia. The school was purposively selected based on its available digital infrastructure (including a one-to-one device ratio in the computer laboratory) and its teachers' prior experience with digital learning tools. Participants consisted of 72 students enrolled in class XI MIPA during the academic year 2023/2024. The sample was divided into two intact classes: XI MIPA 1 served as the control group ($n = 36$) and XI MIPA 2 served as the experimental group ($n = 36$). The age range of participants was 15–17 years. Prior to the intervention, no statistically significant difference was found between the two groups in terms of pretest chemistry scores ($t = 0.23$, $p = 0.819$), confirming initial group equivalence.

AI-powered adaptive learning platform

The AI-powered AI-powered AI-powered adaptive learning platform used in this study was implemented through a customized web-based application designed to deliver chemistry content on three topics: reaction rates, chemical equilibrium, and thermochemistry. The platform employed a Bayesian knowledge tracing algorithm to estimate each student's mastery level for individual knowledge components in real time. Based on these mastery estimates, the system dynamically selected subsequent learning activities from a curated content repository consisting of instructional videos, interactive simulations, worked examples, and formative assessment items. Students in the experimental group accessed the platform for six weeks through scheduled laboratory sessions (90 minutes per session, twice weekly), during which the teacher served as a facilitator and monitor. The control group received teacher-led instruction using the same curricular content delivered through conventional lecture and textbook-based exercises.

Instruments and Data Collection

Two primary instruments were used for data collection. First, a Chemistry Academic Achievement Test (CAAT) consisting of 30 items (20 multiple-choice and 10 short-answer questions) was developed and validated to measure student knowledge of reaction rates, chemical equilibrium, and thermochemistry. The test was reviewed by two chemistry education experts from Universitas Negeri Surabaya and one practising senior high school chemistry teacher. Item analysis yielded an average difficulty index of 0.48 and an average discrimination index of 0.38, both within acceptable ranges. The overall test reliability was $\alpha = 0.84$ (Cronbach's alpha). Second, a Learning Experience Questionnaire (LEQ) comprising 24 statements on a five-point Likert scale was administered to the experimental group following the intervention. The LEQ assessed four dimensions: engagement (6 items), motivation (6 items), perceived usefulness (6 items), and ease of use (6 items). The LEQ achieved a reliability coefficient of $\alpha = 0.87$.

Data Analysis

Academic achievement data were analyzed using independent samples t-tests to compare posttest means between groups, preceded by Shapiro-Wilk normality tests and Levene's homogeneity test. Learning gain was calculated using the normalized gain (N-Gain) formula developed by Hake (1998): $g = (\text{posttest} - \text{pretest}) / (\text{maximum score} - \text{pretest})$. Effect size was quantified using Cohen's d, with $d \geq 0.8$ classified as a large effect. Learning experience data from the LEQ were analyzed descriptively as mean percentage scores, with scores ≥ 80% categorized as 'Very Good.' All inferential analyses were conducted using IBM SPSS Statistics version 26 at a significance level of $\alpha = 0.05$.

RESULTS AND DISCUSSION

Academic Achievement: Pretest and Posttest Comparison

Table 1 presents the descriptive statistics and inferential test results for chemistry academic achievement scores across both groups before and after the intervention.

Table 1. Comparison of Chemistry Academic Achievement Scores

No.	Parameter	Experimental (n=36)	Control (n=36)	t-value	p-value
1	Pretest Mean (SD)	54.8 (9.21)	55.1 (8.97)	0.23	0.819
2	Posttest Mean (SD)	83.7 (7.14)	72.4 (8.63)	6.43	0.000*
3	N-Gain Score (Category)	0.61 (Moderate)	0.35 (Low)	—	—
4	Cohen's d (Effect Size)	1.02 (Large)	—	—	—

*Significant at $p < 0.05$

The pretest comparison confirmed no significant difference between groups prior to intervention ($t = 0.23$, $p = 0.819$), establishing baseline equivalence. Following the six-week intervention, the experimental group's posttest mean (83.7) was substantially higher than the control group's (72.4), a difference that was statistically significant ($t = 6.43$, $p < 0.001$). The N-Gain score of 0.61 for the experimental group falls in the 'Moderate' effectiveness range, while the control group's N-Gain of 0.35 was classified as 'Low.' Cohen's d of 1.02 confirms a large practical effect, indicating that AI-powered adaptive instruction produced meaningful and educationally significant gains beyond those achieved through conventional teaching.

These results can be understood through several interrelated mechanisms. First, the AI-powered AI-powered adaptive learning platform's Bayesian knowledge tracing algorithm ensured that each student received content calibrated to their current mastery level, thereby minimizing cognitive overload for lower-achieving students and preventing stagnation for higher-achieving ones. This dynamic differentiation mirrors the theoretical ideal described by Anderson et al. (1995) in their foundational work on cognitive tutors, where real-time model tracing and mastery-based progression were shown to yield substantially better learning outcomes than paced instruction. Second, the platform's immediate and specific feedback on student responses allowed learners to identify and address misconceptions without waiting for teacher-corrected assignments—a feature that research consistently links to accelerated conceptual development (Koedinger et al., 1997). Third, the self-paced structure of the AI-powered AI-powered adaptive learning platform enabled students to revisit challenging concepts—such as the relationship between activation energy and reaction rate or the mathematical treatment of equilibrium constants—as many times as needed, promoting deeper encoding and longer-term retention.

Academic Achievement by Chemistry Topic

To provide a more granular understanding of learning gains, posttest scores were disaggregated by the three chemistry topics covered in the study. Table 2 presents the mean scores for each topic by group.

Table 2. Posttest Scores by Chemistry Topic

No.	Chemistry Topic	Experimental Mean	Control Mean	t-value	p
1	Reaction Rates	85.3	73.8	5.12	.000*
2	Chemical Equilibrium	82.6	71.2	4.87	.000*
3	Thermochemistry	83.1	72.1	5.03	.000*
	Overall Mean	**83.7**	**72.4**	**6.43**	**.000***

*Significant at $p < 0.05$

As shown in Table 2, the experimental group outperformed the control group on all three chemistry topics, with statistically significant differences in each case (all $p < 0.001$). The most pronounced advantage was observed in Reaction Rates (experimental mean = 85.3 vs. control mean = 73.8), a topic that involves both qualitative understanding of factors affecting reaction speed and quantitative problem-solving using rate laws and the Arrhenius equation. The AI-powered AI-powered adaptive learning platform's ability to present multiple worked examples and branching practice problems—automatically adjusted to match each student's current error patterns—appears to have been particularly beneficial for this topic. The Chemical Equilibrium topic showed the smallest (though still significant) gap, which may reflect the inherent conceptual complexity of dynamic equilibrium that even adaptive pacing cannot fully overcome within a six-week timeframe. These findings suggest that while AI-powered personalization benefits all chemistry topic areas, its impact may vary in magnitude depending on the cognitive demands of individual topics.

Students' Learning Experience

Following the intervention, experimental group students completed the Learning Experience Questionnaire. Table 3 presents the mean percentage scores for each of the four LEQ dimensions.

Table 3. Students' Learning Experience with the AI-powered AI-powered adaptive learning platform (n = 36)

No.	LEQ Dimension	Mean Score (%)	Category
1	Engagement	86.4	Very Good
2	Motivation	84.7	Very Good
3	Perceived Usefulness	89.1	Very Good
4	Ease of Use	82.3	Very Good
	Overall Learning Experience Mean	**85.6**	**Very Good**

The overall learning experience mean of 85.6% places the AI-powered AI-powered adaptive learning platform in the 'Very Good' category. Perceived Usefulness received the highest rating (89.1%), with students frequently reporting in their open-ended responses that the platform helped them identify exactly which concepts they had not yet mastered and provided relevant practice opportunities without having to wade through material they already understood. This echoes findings from Alshammari et al. (2021), who reported that perceived usefulness is the strongest predictor of continued student engagement with digital learning tools, particularly when students can clearly see the connection between platform activities and their academic performance goals.

Engagement scored 86.4%, reflecting students' positive responses to the platform's interactive format, which included drag-and-drop molecular simulations, real-time graphical feedback on progress dashboards, and scenario-based chemistry problems. These features align with the gamification elements identified by Prieto et al. (2019) as effective drivers of student participation in digital science learning environments. Motivation received a score of 84.7%, indicating that the adaptive system's provision of achievable short-term challenges—calibrated just above each student's current proficiency level—was effective in sustaining intrinsic motivation over the six-week period. This finding supports the theoretical predictions of flow theory (Csikszentmihalyi, 1990, as cited in Mayer, 2019), which posits that optimal engagement occurs when task challenge is appropriately matched to learner skill level.

The lowest-rated dimension was Ease of Use (82.3%), still within the 'Very Good' range, though students' comments highlighted occasional technical difficulties including slow page-loading speeds during simultaneous multi-user sessions and a non-intuitive navigation menu in the simulation module. These technical limitations point to areas for platform refinement and are consistent with findings from Dziuban et al. (2018) and Luan et al. (2020), who noted that the infrastructural and interface quality of AI-driven learning tools significantly moderates the extent of their educational benefit, particularly in school contexts where students may have limited prior experience with sophisticated digital environments. Future iterations of the platform should prioritize usability testing with secondary school student populations to ensure that interface complexity does not serve as a barrier to learning.

Implications for Chemistry Education

The results of this study carry meaningful implications for chemistry educators and educational technology developers operating in the Indonesian secondary school context. First, the large effect size ($d = 1.02$) obtained in this study suggests that AI-powered AI-powered adaptive learning platforms can generate gains that substantially exceed those typically observed with conventional instruction—even in a relatively short six-week implementation window. This is particularly significant for topics such as reaction rates and chemical equilibrium, which have long been identified as conceptual stumbling blocks for Indonesian senior high school students. Second, the high perceived usefulness scores indicate that students are capable of recognizing and articulating the learning benefits of personalized digital instruction, which has positive implications for adoption and sustained use. Third, the ease of use concerns raised by students underscore the importance of ensuring reliable technical infrastructure and conducting iterative usability testing before large-scale deployment of AI platforms in schools.

Several limitations should be acknowledged when interpreting the findings of this study. First, the research was conducted in only one public senior high school in Surabaya, which means that the findings should be interpreted as context-specific evidence rather than as broadly representative of all Indonesian senior high schools. Differences in school infrastructure, teacher readiness, student characteristics, and local instructional practices across regions may influence the effectiveness of the AI-powered AI-powered adaptive learning platform. Second, the sample size was relatively modest and limited to 72 Grade XI science-track students, which further restricts the generalizability of the results to other grade levels, streams, subject areas, or school types. Therefore, caution is needed in extending these findings beyond the specific context of this study. Future research involving multiple schools, larger samples, and more diverse educational settings is necessary to strengthen the external validity of the findings.

CONCLUSION

This study demonstrated that an AI-powered AI-powered adaptive learning platform significantly improved chemistry academic achievement among class XI MIPA students at SMA Negeri 3 Surabaya. The experimental group achieved a substantially higher posttest mean (83.7) compared to the control group (72.4), with a large effect size (Cohen's d = 1.02) and an N-Gain score of 0.61 (Moderate), versus 0.35 (Low) for the control group. Significant academic advantages were observed across all three chemistry topics—reaction rates, chemical equilibrium, and thermochemistry. Students' overall learning experience with the AI-powered AI-powered adaptive learning platform was rated positively (mean = 85.6%, Very Good), with perceived usefulness receiving the highest score and ease of use the lowest, pointing to the need for continued interface refinement. These findings affirm that AI-driven personalization is a promising strategy for addressing the persistent challenge of academic heterogeneity in secondary chemistry classrooms and for fostering deeper conceptual understanding through individually calibrated learning trajectories. Future research should examine the sustainability of these learning gains over longer periods, explore the platform's applicability to other chemistry topics and grade levels, and investigate the moderating role of teacher facilitation quality on AI-powered AI-powered adaptive learning platform effectiveness in Indonesian school settings.

REFERENCES

Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education, 32*(7), 939–961. https://doi.org/10.1080/09500690902890130

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*(1), 66–71. https://doi.org/10.1119/1.12989

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Dagher, Z. R. (1994). Does the use of analogies contribute to conceptual change? *Science Education, 78*(6), 601–614. https://doi.org/10.1002/sce.3730780605

Docktor, J. L., & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. *Physical Review Special Topics - Physics Education Research, 10*(2), 020119. https://doi.org/10.1103/PhysRevSTPER.10.020119

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education, 25*(6), 671–688. https://doi.org/10.1080/09500690305016

Duit, R., & Treagust, D. F. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education, 3*(2), 297–328. https://doi.org/10.1007/s11422-008-9090-4

Eryilmaz, A. (2002). Effects of conceptual assignments and conceptual change discussions on students' misconceptions and achievement regarding force and motion. *Journal of Research in Science Teaching, 39*(10), 1001–1015. https://doi.org/10.1002/tea.10054

Fetherstonhaugh, T., & Treagust, D. F. (1992). Students' understanding of light and its properties: Teaching to engender conceptual change. *Science Education, 76*(6), 653–672. https://doi.org/10.1002/sce.3730760606

Graham, T., Berry, J., & Rowlands, S. (2013). Are 'misconceptions' or alternative frameworks of force and motion spontaneous or formed prior to instruction? *International Journal of Mathematical Education in Science and Technology, 44*(1), 84–103. https://doi.org/10.1080/0020739X.2012.703333

Halloun, I. A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics, 53*(11), 1056–1065. https://doi.org/10.1119/1.14031

Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics, 53*(11), 1043–1055. https://doi.org/10.1119/1.14030

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*(3), 141–158. https://doi.org/10.1119/1.2343497

Hynd, C. R. (2001). Refutational texts and the change process. *International Journal of Educational Research, 35*(7–8), 699–714. https://doi.org/10.1016/S0883-0355(02)00010-1

Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2016). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education, 12*(2), 341–374. https://doi.org/10.12973/eurasia.2015.1369a

Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education, 35*(2), 238–260. https://doi.org/10.1080/02635143.2017.1310094

Limon, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*(4–5), 357–380. https://doi.org/10.1016/S0959-4752(00)00037-2

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Lawrence Erlbaum Associates.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227. https://doi.org/10.1002/sce.3730660207

Suparno, P. (2013). *Miskonsepsi dan perubahan konsep dalam pendidikan fisika*. PT Grasindo.

Terry, C., & Jones, G. (1986). Alternative frameworks: Newton's third law and conceptual change. *European Journal of Science Education, 8*(3), 291–298. https://doi.org/10.1080/0140528860080305

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics, 66*(4), 338–352. https://doi.org/10.1119/1.18863

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education, 10*(2), 159–169. https://doi.org/10.1080/0950069880100204

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*(1), 45–69. https://doi.org/10.1016/0959-4752(94)90018-3

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