pISSN 2685-0761
eISSN 2685-0850

JURNAL INOVASI PEMBELAJARAN KIMIA
(Journal of Innovation in Chemistry Education)
https://jurnal.unimed.ac.id/2012/index.php/jipk
email: Jinovpkim@unimed.ac.id

Recieved : 29 August 2026
Revised : 05 September 2026
Accepted : 05 October 2026
Published : 31 October 2026
Pages : 299 – 316

Designing Diagnosis-Informed Microlearning with AI-Feedback for Stoichiometry Learning

Jam'an Fahmi[1]*, Muhammad Isa Siregar[2], Hafiz Irwandi[3], Aralia Putri Marsyah[4], Nazwa Azhara[5], Hilma Tsabita Hasibuan[6], Nazwa Azzahra[7], and Adelia Ofira Pakpahan[8]

[1,2,3,4,5,6,7,8]*Chemistry Education Study Program, Universitas Negeri Medan, Medan*
**Email:* jamanfahmi@unimed.ac.id

Abstract: *Stoichiometry is conceptually challenging because students may experience misconceptions that are not adequately revealed through conventional achievement tests. This study aimed to design and conduct a limited proof-of-concept of diagnosis-informed microlearning with AI-feedback to support students' conceptual understanding of stoichiometry. The study involved 68 Grade X students from two senior high schools and five expert validators. Thirty diagnostic items were validated, with 24 meeting the predetermined content-validity criteria, from which 18 items were selected to represent six stoichiometry subtopics. A separate 30-item conceptual understanding test yielded 24 valid items, from which 20 items were selected for limited implementation. The initial diagnostic assessment was administered to 34 students, while the revised learning website was implemented with another 34 students. The diagnostic results showed the highest proportion of conceptual problems in limiting reagents (67.7%). Following the limited implementation, 52.94% of students were classified as having moderate-to-high conceptual understanding. Students' responses to the AI-feedback website were generally positive, with positive responses exceeding 75% across all indicators. The findings provide initial evidence of the potential of diagnosis-informed AI-feedback for formative stoichiometry learning, while further controlled studies are required to establish its effectiveness.*

Keywords: *Stoichiometry; misconception; diagnostic assessment; microlearning; AI-feedback*

INTRODUCTION

Stoichiometry is a fundamental component of chemistry learning because it provides the quantitative basis for interpreting chemical reactions and connects concepts such as moles, chemical equations, stoichiometric ratios, and quantities of substances. However, stoichiometry remains challenging because students must coordinate conceptual understanding with symbolic representations, quantitative relationships, and algorithmic procedures (Fitriani et al., 2022; Krause et al., 2026; Stott, 2023). Difficulties may also be associated with how stoichiometric concepts are represented and sequenced in learning resources (Karimatunnisa' & Aloysius, 2023). Consequently, conventional assessments that

emphasize final numerical answers may not adequately reveal the conceptual sources of students' difficulties.

Misconceptions may persist even when students are able to perform routine calculations. Studies have identified stoichiometric misconceptions involving mole relationships, limiting reagents, and chemical formulas (Aini et al., 2016; Lestari et al., 2021; Sa'adah et al., 2022). Diagnostic approaches, particularly three-tier instruments that combine answer selection, reasoning, and confidence, can distinguish misconceptions from other forms of incorrect or uncertain responses (Ekawisudawati et al., 2021; Hasniyah & Muchtar, 2021). Recent work has also demonstrated the use of three-tier diagnostic assessment for identifying stoichiometric misconceptions (Gani, 2025), while digital learning has shown potential for supporting misconception detection and remediation (Alfiyanti & Sukarmin, 2021). Thus, diagnostic assessment can provide more instructionally relevant information than achievement scores alone by indicating not only whether students are incorrect but also where conceptual difficulties occur.

The instructional value of diagnosis, however, depends on how the resulting information is used to support subsequent learning. Formative feedback can guide students' next learning actions, but providing timely and individualized responses to different conceptual difficulties remains challenging. Digital chemistry learning environments have demonstrated the feasibility of automated, individualized, error-specific, and just-in-time feedback (Eitemüller et al., 2023), while broader evidence indicates that explanatory feedback in technology-rich environments can support learning achievement (Cai et al., 2023). Existing chemistry learning media, including website-, Android-, and electronic-resource-based environments, have likewise demonstrated the feasibility of digital learning support (Arief et al., 2021; Rilanty & Juwitaningsih, 2020; Panggabean & Purba, 2021; Silaban & Sianturi, 2021; Silaban & Panggabean, 2022). However, these approaches have generally emphasized media development, feasibility, student responses, or learning outcomes rather than using students' diagnosed conceptual difficulties to determine subsequent individualized feedback.

Microlearning provides a suitable structure for connecting conceptual difficulties with focused learning resources because content can be organized into short units directed toward specific learning objectives. Research has emphasized the importance of clear objectives, appropriately sized content, learner engagement, and suitable delivery modes in microlearning (Alias & Abdul Razak, 2023), while recent evidence has associated microlearning with cognitive, behavioural, and affective outcomes and reported improved learning performance (Balasundaram et al., 2024; Monib et al., 2025). This structure is particularly relevant to stoichiometry because its interconnected concepts can be organized into focused learning units.

Artificial intelligence (AI) further expands the possibility of delivering individualized feedback at scale. Large language models have demonstrated the ability to generate readable and relatively consistent assessment feedback, with GPT-4 showing stronger performance than earlier models and human instructors on several feedback dimensions (Dai et al., 2024). LLM-generated feedback has also been associated with students' revision performance and motivational responses (Meyer et al., 2024), while meta-analytic evidence indicates that human–machine feedback can positively influence learning processes and outcomes, although its effects vary according to the form and direction of feedback (Liao et al., 2023). Nevertheless, technically fluent AI-generated feedback does not necessarily address the conceptual source of an error; personalized feedback research emphasizes the importance of the information used to adapt feedback to individual learners (Maier & Klotz, 2022).

These findings reveal a gap between diagnosing conceptual difficulties and using diagnostic information to determine subsequent individualized feedback within a digital learning environment. Existing studies have established the value of diagnostic assessment, digital chemistry learning, microlearning, and automated or AI-supported feedback, but these components have not been sufficiently integrated into a learning sequence in which students' diagnosed conceptual difficulties directly inform the focus of subsequent AI-feedback. The present study addresses this gap by integrating diagnostic assessment, focused microlearning, and diagnosis-informed AI-supported formative feedback for stoichiometry learning. Rather than providing identical explanations to all students who respond incorrectly, the proposed environment uses students' diagnosed conceptual difficulties to determine the direction and focus of subsequent feedback.

Accordingly, this study aims to design and conduct a limited proof-of-concept implementation of microlearning-based chemistry learning with AI-feedback informed by students' diagnosed conceptual difficulties in stoichiometry. The study examines the conceptual difficulties identified through diagnostic assessment, the implementation of diagnosis-informed AI-feedback, students' conceptual understanding following the limited implementation, and their responses to the AI-feedback. The study is positioned as an early-stage investigation that provides initial evidence regarding the feasibility and potential of this integrated approach rather than definitive evidence that AI-feedback eliminates students' misconceptions.

LITERATURE REVIEW

A. Stoichiometric Conceptual Difficulties and Diagnostic Assessment

Stoichiometric competence involves interconnected conceptual and procedural knowledge rather than calculation accuracy alone (Krause et al., 2026; Stott, 2023). Students may therefore produce similar incorrect answers for different conceptual reasons. Multiple-tier diagnostic assessment can address this problem by considering students' answers together with their reasoning and confidence. Such instruments have been widely used to identify misconceptions in science and chemistry learning (Soeharto et al., 2019; Mubarokah et al., 2018; Shiddiqi et al., 2024). In stoichiometry, diagnostic approaches have identified difficulties involving mole relationships, limiting reagents, and empirical and molecular formulas (Aini et al., 2016; Lestari et al., 2021). Accordingly, diagnostic assessment provides information about the conceptual location and nature of students' difficulties that cannot be obtained from final achievement scores alone.

B. Microlearning and Digital Chemistry Learning

Digital chemistry learning environments have demonstrated the feasibility of website-, Android-, and electronic-resource-based learning support (Rilanty & Juwitaningsih, 2020; Panggabean & Purba, 2021; Silaban & Sianturi, 2021; Silaban & Panggabean, 2022). However, these studies have primarily emphasized media development, feasibility, practicality, student responses, or learning outcomes rather than using diagnostic information to determine individualized learning support. Microlearning offers a more focused structure by organizing content into short units directed toward specific learning objectives. Effective microlearning requires clear objectives, appropriately sized content, learner engagement, and suitable delivery modes (Alias & Abdul Razak, 2023). Recent evidence further indicates positive cognitive, behavioural, and affective outcomes associated with microlearning and improved learning performance following microlearning interventions (Balasundaram et

al., 2024; Monib et al., 2025). These characteristics make microlearning relevant for organizing stoichiometry into focused conceptual units.

C. Automated and AI-Supported Formative Feedback

Automated feedback can provide timely and individualized responses, with digital chemistry exercises demonstrating the potential for error-specific and just-in-time formative feedback (Deeva et al., 2021; Eitemüller et al., 2023). Evidence from technology-rich learning environments also indicates that explanatory feedback can support learning achievement (Cai et al., 2023). Generative AI extends this potential by enabling feedback generation at scale. Dai et al. (2024) reported that large language models could generate readable and relatively consistent assessment feedback, while Maier and Klotz (2022) emphasized that effective personalization depends on the information used to adapt feedback to individual learners. LLM-generated feedback has also been associated with revision performance and motivational responses (Meyer et al., 2024), and meta-analytic evidence suggests positive effects of human–machine feedback, although outcomes vary according to feedback characteristics (Liao et al., 2023). These findings indicate that AI-feedback should not be viewed simply as automated explanation generation; its instructional value depends on whether the feedback is grounded in relevant evidence about students' learning difficulties.

D. Research Position

Taken together, previous studies establish the value of diagnostic assessment for identifying conceptual difficulties (Soeharto et al., 2019; Mubarokah et al., 2018; Aini et al., 2016; Lestari et al., 2021), digital chemistry learning and microlearning for organizing focused learning resources (Rilanty & Juwitaningsih, 2020; Alias & Abdul Razak, 2023; Monib et al., 2025), and automated or AI-supported feedback for providing timely and individualized learning support (Deeva et al., 2021; Eitemüller et al., 2023; Dai et al., 2024; Meyer et al., 2024). However, these components have generally been examined separately. The present study connects them through a diagnosis-informed sequence in which students' identified conceptual difficulties guide the focus of subsequent microlearning and differentiated AI-feedback. Thus, the contribution lies in linking diagnostic evidence with targeted formative feedback rather than in the use of AI or microlearning alone.

METHODS

A. Research Design and Participants

This study employed an early-stage R&D design with a limited-scale proof-of-concept approach. The study comprised five stages: instrument development and expert validation, initial diagnostic assessment, misconception mapping and diagnosis-informed microlearning/AI-feedback design, limited website implementation, and post-implementation assessment. Existing open-access YouTube videos were curated as short microlearning resources and aligned with the identified conceptual difficulties. The study was descriptive and was not designed to establish causal effectiveness.

The participants were 68 Grade XI students from two senior high schools in Deli Serdang, North Sumatra, Indonesia (34 students per school). The private-school group participated in the initial diagnostic assessment, whereas the public-school group participated in the limited implementation. The groups served sequential rather than experimental-control purposes.

B. Instrument Development, Validation, and Diagnosis-Informed Learning Design

Two 30-item instruments were initially developed: a three-tier diagnostic test to identify students' conceptual difficulties in stoichiometry and a conceptual understanding test to assess students' understanding following the limited implementation. The diagnostic instrument represented six stoichiometry subtopics: mole concept,

balancing chemical equations, limiting reagents, empirical and molecular formulas, hydrated compounds, and percentage yield. The items were developed according to the corresponding learning indicators. The three-tier structure consisted of an answer tier, a reason tier, and a confidence tier, allowing students' conceptual difficulties to be interpreted beyond whether their selected answer was correct or incorrect.

Content validation was conducted by five expert lecturers with relevant expertise in chemistry education, chemistry content, assessment, and educational media. The experts evaluated item relevance, conceptual accuracy, clarity, and suitability. Content validity was quantified using the Content Validity Ratio (CVR) proposed by Lawshe (1975) and the Content Validity Index (CVI) approach described by Polit et al. (2007).

The CVR was calculated as: $CVR = \frac{n_{e\,-N/2}}{N/2}$

where n_e is the number of experts judging an item as essential and N is the total number of experts.

The I-CVI was calculated as: $CVI = \frac{n_v}{N}$

where *nv* is the number of experts assigning the item a relevant or valid rating and N is the total number of experts.

For the diagnostic instrument, 22 of the initial 30 items met the CVR and CVI requirements. Eighteen items were subsequently selected to represent the six stoichiometry subtopics, with three items representing each subtopic. For the conceptual understanding instrument, 24 of the initial 30 items met the content-validity criteria, from which 20 items were selected for the limited implementation. The selected items were revised based on expert recommendations before administration.

The 18-item three-tier diagnostic test was administered to 34 students from the private school. Students' answer, reason, and confidence responses were used to classify conceptual difficulties and identify misconception patterns. The classification was subsequently linked to the question indicator and corresponding stoichiometry subtopic represented by each item. Because each subtopic was represented by three diagnostic items, interpretation was restricted to the conceptual indicators represented by the selected items rather than generalized to all possible misconceptions within the subtopic.

The resulting misconception map was then used to design the subsequent microlearning and AI-feedback. Existing instructional videos available on YouTube were curated and organized into short learning resources according to the classroom learning sequence and the conceptual difficulties identified through the diagnostic test. The AI-feedback was designed as a formative support mechanism in which the focus of feedback was determined by the conceptual difficulty indicated by students' responses. Thus, students with different diagnosed difficulties could receive different explanations, conceptual prompts, or recommendations for relevant microlearning resources rather than a uniform response to an incorrect answer. This established the intended sequence of diagnosis → conceptual difficulty → focused microlearning → differentiated AI-feedback.

C. Limited Implementation

The revised AI-feedback website was implemented on a limited scale with the 34 students from the public school. Students accessed the microlearning materials and completed the stoichiometry assessment activities through the website.

The learning sequence connected the diagnostic information, selected microlearning resources, and AI-feedback. The YouTube videos functioned as supporting learning resources and were arranged according to the classroom learning sequence. The AI-feedback provided additional conceptual guidance based on students' responses.

a. Conceptual Understanding Assessment

Following the limited implementation, the 34 students completed a 20-item conceptual understanding test. The 20 items were selected from the 24 items that had met the predetermined CVR and CVI criteria from the original 30-item instrument.

The students' scores were converted to a 0–100 scale using:

$$S = \frac{R}{N} \text{ x } 100$$

where S is the conceptual understanding score, R is the number of correct responses, and N is the total number of test items.

The scores were categorized as follows:

Score Range	Conceptual Understanding
0–54	Very Low
55–69	Low
70–84	Moderate
85–100	High

The resulting distribution was used to describe the conceptual understanding profile of students participating in the limited implementation.

b. Students' Responses to AI-Feedback

Students' perceptions of the AI-feedback were collected using a response questionnaire administered after the website implementation. The questionnaire examined students' perceptions of the **acceptance, clarity, usefulness, and assistance provided by the AI-feedback**.

The percentage of positive responses was calculated as:

$$P = \frac{f}{N} x\ 100\%$$

where P is the percentage of positive responses, f is the number of students giving a positive response, and N is the total number of respondents.

The results were interpreted using four categories:

Positive Response	Interpretation
0–25%	Very Poor
26–50%	Poor
51–75%	Good
76–100%	Very Good

The analysis was used to describe the extent to which students perceived the AI-feedback as understandable and helpful during the limited implementation.

c. Data Analysis

The study primarily used descriptive quantitative analysis. Expert-validation data were analyzed using CVR and CVI. Diagnostic-test responses were analyzed to determine the distribution of misconceptions and lack of conceptual understanding across the six stoichiometry subtopics, while the initial student responses were also used for quantitative item analysis.

For the limited implementation group, conceptual understanding was analyzed using the 20-item test and summarized through score distributions and categorical frequencies. Students' responses to the AI-feedback questionnaire were analyzed using percentages.

RESULT AND DISCUSSION

A. Initial Diagnostic Results and Misconception Mapping

The initial diagnostic assessment was conducted to identify students' conceptual difficulties in stoichiometry before the implementation of the AI-feedback-assisted learning website. Eighteen diagnostic items representing six stoichiometry subtopics were administered to 34 students. Each subtopic was represented by three items, allowing the students' responses to be examined in a balanced manner across the targeted conceptual areas.

The diagnostic results were classified into two major categories of conceptual problems, namely misconception and lack of conceptual understanding. The percentage of students in each category was calculated for each subtopic. The combined percentage of these two categories was used to describe the overall proportion of students experiencing conceptual problems.

Table 1. Distribution of Students' Conceptual Problems Based on the Initial Diagnostic Test

Stoichiometry Subtopic	Misconception (%)	Lack of Conceptual Understanding (%)	Total Conceptual Problems (%)	Interpretation
Balancing Chemical Equations	35.3	15.7	**51.0**	Needs reinforcement
Mole Concept	22.5	24.5	**47.0**	Needs reinforcement
Limiting Reagents	32.4	35.3	**67.7**	Very high priority
Percent Yield	12.7	28.4	**41.1**	Relatively better
Empirical and Molecular Formulas	36.3	18.6	**54.9**	Very high priority
Hydrated Compounds	15.7	31.4	**47.1**	Needs reinforcement

Conceptual problems occurred across all six subtopics, with limiting reagents showing the highest overall proportion (67.7%), followed by empirical and molecular formulas (54.9%) and balancing equations (51.0%). Empirical and molecular formulas had the highest misconception proportion (36.3%), whereas limiting reagents showed the largest combined difficulty. These findings indicate that conceptual problems varied across subtopics and that misconception and lack of understanding require different forms of formative support.

B. Misconception Patterns Across Stoichiometry Subtopics

The diagnostic results were further examined to identify the conceptual patterns underlying students' misconceptions across the six stoichiometry subtopics. Because each subtopic was represented by three diagnostic items, the interpretation was restricted to misconceptions that were conceptually relevant to the indicators measured by those items. Previous studies were used as supporting references to contextualize the identified patterns rather than to assume that all previously reported misconceptions were present among the participants.

Table 2. Mapping of Diagnostic Indicators and Relevant Misconception Patterns in Stoichiometry

Subtopic	Misconception (%)	Relevant Misconception Patterns
Balancing Chemical Equations	**35.3**	(1) Treating subscripts as quantities that can be changed to balance an equation; (2) interpreting stoichiometric coefficients as mass relationships rather than mole relationships; (3) considering an equation balanced when the number of molecules on both sides is the same.
Mole Concept	**22.5**	(1) Assuming that one mole always corresponds to one gram; (2) treating moles and mass as directly interchangeable without considering molar mass; (3) assuming that 22.4 L is the molar volume of a gas under all temperature and pressure conditions.
Limiting Reagents	**32.4**	(1) Identifying the limiting reagent solely as the reactant with the smallest mass; (2) assuming that the excess reactant is completely consumed; (3) assuming that the first reactant written in an equation is always the limiting reagent.
Percent Yield	**12.7**	(1) Assuming that a yield greater than 100% can occur under ideal conditions; (2)

Subtopic	Misconception (%)	Relevant Misconception Patterns
		determining theoretical yield from all reactants rather than from the limiting reagent; (3) interpreting a low yield as necessarily indicating that the product is impure.
Empirical and Molecular Formulas	**36.3**	(1) Assuming that an empirical formula is always identical to a molecular formula; (2) interpreting subscripts as representations of atomic mass rather than the number or ratio of atoms; (3) assuming that identical percentage compositions necessarily indicate identical compounds.
Hydrated Compounds	**15.7**	(1) Treating the dot in a hydrate formula merely as a separator rather than as part of a stoichiometric relationship; (2) assuming that crystal water can be removed through cooling; (3) calculating the percentage of water using the mass of the anhydrous salt rather than the total mass of the hydrate.

The pattern analysis reveals that the most prominent misconceptions were found in empirical and molecular formulas (36.3%), balancing chemical equations (35.3%), and limiting reagents (32.4%). These findings complement the results in Table 1, where limiting reagents had the highest overall proportion of conceptual problems. The distinction occurs because the total conceptual-problem category includes both misconception and lack of conceptual understanding. Thus, a subtopic may have a high overall difficulty while exhibiting a comparatively lower proportion of established misconceptions.

For balancing chemical equations, the identified patterns point to difficulties in distinguishing the role of coefficients and subscripts. This distinction is fundamental because changing a subscript changes the identity or composition of a chemical substance, whereas changing a coefficient changes the stoichiometric amount of the substance. The finding therefore suggests that feedback for this concept should not simply display the correctly balanced equation, but should emphasize the conceptual meaning of coefficients and the conservation of atoms.

In the mole concept, the misconception percentage was lower (22.5%), but the patterns indicate difficulties in connecting the mole with other chemical quantities. The misconceptions related to equating one mole with one gram or treating mass and amount of substance as directly interchangeable suggest that students may rely on numerical associations rather than understanding the role of molar mass as a conversion factor. Consequently, microlearning and AI-feedback for this topic should emphasize the relationships among amount of substance, mass, number of particles, and gas volume.

The limiting reagent results are particularly important because this subtopic produced the highest overall conceptual problem (67.7%) while also showing a relatively high misconception proportion (32.4%). The identified patterns suggest that students may use superficial cues, such as choosing the reactant with the smallest mass, rather than comparing the available amount of each reactant with the stoichiometric requirements of the reaction. This provides a clear rationale for targeted feedback that guides students through the sequence of mass → mole → stoichiometric ratio → limiting reagent, rather than immediately providing the final answer.

For empirical and molecular formulas, the misconception proportion was the highest among the six subtopics (36.3%). The patterns indicate possible confusion between empirical and molecular representations and between numerical subscripts and atomic quantities. This finding is pedagogically important because solving these problems requires students to coordinate chemical composition, mole ratios, and molecular-scale representation. AI-feedback for this concept

should therefore emphasize the conceptual relationship between percentage composition, mole ratio, empirical formula, and molecular formula.

The lower misconception proportions observed for hydrated compounds (15.7%) **and** percent yield (12.7%) do not indicate that these topics are free from conceptual difficulties. Rather, Table 1 shows that a substantial proportion of students in these topics were categorized as lacking conceptual understanding. Therefore, the distinction between *misconception* and *lack of understanding* remains important in determining the type of formative support required. Students who lack understanding may need foundational explanations, whereas students who hold misconceptions may require feedback that explicitly challenges and restructures their existing conception.

C. Implications of Diagnostic Findings for AI-Feedback Design

The diagnostic findings provided an empirical basis for determining the conceptual focus of the microlearning and AI-feedback developed in this study. Rather than delivering the same feedback to all students, the identified conceptual difficulties were used to determine the concepts requiring greater instructional emphasis. The design therefore followed a diagnosis-informed formative feedback approach, in which the students' initial conceptual difficulties became a reference for selecting learning resources and formulating feedback.

The diagnostic results indicated that limiting reagents, empirical and molecular formulas, and balancing chemical equations required particular attention. Limiting reagents showed the highest proportion of overall conceptual problems (67.7%), whereas empirical and molecular formulas showed the highest proportion of misconceptions (36.3%). These findings were therefore considered when determining the conceptual emphasis of the learning resources and feedback.

Table 3. Alignment of Diagnostic Findings with Microlearning and AI-Feedback Design

Stoichiometry Subtopic	Diagnostic Finding	Priority Conceptual Focus	Microlearning Focus	AI-Feedback Focus
Balancing Chemical Equations	35.3% misconception; 51.0% total conceptual problems	Conservation of mass, reaction coefficients, and distinction between coefficients and subscripts	Balancing equations using conservation of atoms and appropriate coefficients	Guide students to distinguish coefficients from subscripts and reconsider the relationship between coefficients and amounts of substances
Mole Concept	22.5% misconception; 47.0% total conceptual problems	Relationships among mole, mass, particles, and gas volume	Conversion among amount of substance, mass, particles, and gas volume	Reinforce the role of molar mass and Avogadro's number in quantitative conversions
Limiting Reagents	32.4% misconception; 67.7% total conceptual problems	Mole ratio, available amount of reactants, and identification of the limiting reagent	Determining limiting reagents through mole-based stoichiometric comparisons	Direct students from mass to mole and stoichiometric ratio before determining the limiting reagent

Stoichiometry Subtopic	Diagnostic Finding	Priority Conceptual Focus	Microlearning Focus	AI-Feedback Focus
Percent Yield	12.7% misconception; 41.1% total conceptual problems	Theoretical yield, actual yield, and meaning of percent yield	Relationship between theoretical and experimental yield	Clarify the distinction between theoretical and actual yield and the basis for calculating percent yield
Empirical and Molecular Formulas	36.3% misconception; 54.9% total conceptual problems	Mole ratios, empirical formula, molecular formula, and Mr	Determining empirical and molecular formulas from composition data	Help students distinguish empirical from molecular formulas and interpret subscripts as atom ratios
Hydrated Compounds	15.7% misconception; 47.1% total conceptual problems	Crystal water and stoichiometric relationship between hydrated and anhydrous compounds	Concept and calculation of water of crystallization	Reinforce the stoichiometric relationship between the hydrate and its water content

The microlearning component was not developed as a set of newly produced instructional videos. Instead, existing instructional videos available on YouTube were selected and curated according to the identified conceptual needs and aligned with the learning sequence implemented in the classroom**.** This approach allowed the learning resources to remain consistent with the actual instructional timeline while providing short, focused learning materials corresponding to the concepts requiring reinforcement.

The AI-feedback was subsequently positioned as a formative component accompanying the learning and assessment activities. The feedback was designed to provide conceptual guidance rather than simply indicate whether an answer was correct or incorrect. For example, difficulties related to limiting reagents were connected to the relationship between mass, moles, stoichiometric ratios, and the amount of product, while difficulties related to empirical and molecular formulas were connected to the relationship between elemental composition, mole ratios, and molecular representation.

The alignment between diagnostic findings and feedback design is important because the purpose of the AI-feedback in this study was not to claim that students' misconceptions had been eliminated, but to demonstrate how diagnostic information could be translated into targeted formative support. In this sense, the diagnostic assessment functioned as an upstream component of the feedback system: the identified conceptual difficulties informed the selection of learning resources and the conceptual direction of the feedback provided through the website.

This design also distinguishes the present approach from feedback that merely provides the correct answer. A student who selects an incorrect answer may require different support depending on whether the response reflects a lack of understanding or an alternative conception. Therefore, the diagnostic information was used to establish a more conceptually targeted feedback structure.

The differentiated nature of the AI-feedback represents an important feature of the present design. Unlike generic AI feedback that may provide a similar explanation for different incorrect responses, the feedback in this study was informed by the

conceptual difficulty identified from students' diagnostic responses. This design is consistent with previous research showing that adaptive feedback can be generated in response to learners' actions or performance data (Sailer et al., 2023; Colliot et al., 2024; Huang et al., 2025). Research on digital adaptive learning has also emphasized that feedback should be tailored to learners' individual needs rather than delivered uniformly (Mejeh et al., 2024). In chemistry education, Pölloth (2025) demonstrated the feasibility of adaptive feedback in an online chemistry learning environment. The present design extends this direction by using students' diagnosed conceptual difficulties as the basis for determining the focus of AI-generated formative feedback. Thus, the distinguishing feature is not simply the use of generative AI, but the connection between diagnostic evidence and differentiated feedback.

This distinction should nevertheless be interpreted cautiously. Adaptive AI-feedback does not inherently guarantee superior learning outcomes. Bauer et al. (2025), for example, found no significant performance advantage of AI-generated adaptive feedback over static expert feedback in a highly structured statistical learning context. The effects of adaptive feedback may depend on task characteristics, the information used for adaptation, and the quality and relevance of the feedback itself. Therefore, the present study does not claim that differentiated AI-feedback is inherently more effective than generic feedback. Instead, the limited implementation provides initial evidence that diagnostic information can be operationalized to produce feedback that is conceptually differentiated according to students' identified difficulties. This diagnosis-informed mechanism provides a basis for further controlled investigation of whether such differentiation produces measurable advantages in conceptual learning.

D. Students' Conceptual Understanding after the Limited Implementation

Following the implementation of the microlearning materials and AI-feedback-assisted assessment website, the 34 students in the public school completed a 20-item conceptual understanding test. The test items were selected from the 24 items that had met the predetermined expert-validation criteria from the initial 30-item instrument. The assessment was intended to provide preliminary evidence of students' conceptual understanding after interacting with the developed learning environment.

Students' scores were classified into four categories: high (85–100), moderate (70–84), low (55–69), and very low (0–54). The distribution of students across these categories is presented in Table 4.

Table 4. Distribution of Students' Conceptual Understanding after the Limited Implementation

Conceptual Understanding Category	Score Range	Number of Students	Percentage
High	85–100	3	8.82%
Moderate	70–84	15	44.12%
Low	55–69	4	11.76%
Very Low	0–54	12	35.29%
Total	—	**34**	**100%**

The distribution shows that the largest proportion of students (15 students; 44.12%) was in the moderate conceptual understanding category. Three students (8.82%) achieved the high category, while four students (11.76%) were categorized as having low conceptual understanding. Meanwhile, 12 students (35.29%) were in the very low category.

Overall, 18 students (52.94%) were classified in the moderate-to-high categories, whereas 16 students (47.06%) remained in the low-to-very-low categories. This distribution indicates that students' conceptual understanding was heterogeneous after the limited implementation. The presence of a

relatively large moderate group suggests that the learning environment provided a basis for conceptual engagement, while the substantial low and very-low groups indicate that students did not respond uniformly to the learning resources and AI-feedback.

This finding is particularly relevant when considered together with the initial diagnostic results. The diagnostic assessment had previously shown that conceptual difficulties varied across stoichiometry subtopics, with limiting reagents, empirical and molecular formulas, and balancing chemical equations emerging as important areas requiring reinforcement. Therefore, the variation in conceptual understanding observed in the limited implementation reinforces the importance of differentiated and diagnosis-informed formative support rather than relying on a uniform learning treatment.

However, the present result should not be interpreted as evidence that AI-feedback improved or eliminated students' misconceptions**.** The study employed a limited-scale proof-of-concept design without a control group or a comparable pretest-posttest measure using the same conceptual-understanding instrument. Consequently, the distribution in Table 4 is best interpreted as initial evidence of students' conceptual understanding following interaction with the developed website, rather than as a measure of causal learning gains. These result demonstrates heterogeneity in students' conceptual understanding following the limited implementation and highlights the need for more differentiated feedback based on students' diagnosed conceptual difficulties.

E. Students' Responses to the AI-Feedback Website

Following the limited implementation, students were asked to provide responses regarding their experience using the AI-feedback-assisted website. The questionnaire consisted of 15 statements covering five aspects: ease of access, learning efficiency, media attractiveness and interactivity, perceived benefits of AI-feedback, and learning motivation. Each statement was responded to using a four-point Likert scale ranging from strongly agree to strongly disagree.

The analysis in this study was focused on the proportion of students providing positive responses**,** represented by the combination of the strongly agree and agree categories. The results are presented in Table 5.

Table 5. Students' Positive Responses to the AI-Feedback Website

Aspect	Statement	Positive Response (f)	Percentage
Ease of Access	The AI-feedback website is easy to access using a laptop or mobile phone.	34	**100.00%**
Ease of Access	I can understand how to use the menus and features of the website without difficulty.	34	**100.00%**
Ease of Access	The dashboard, question navigation, and buttons on the website are easy to understand.	32	**94.12%**
Ease of Access	The login process, opening the test, and answering questions can be completed easily.	34	**100.00%**
Learning Efficiency	The website helps me complete the test more practically and systematically.	34	**100.00%**
Learning Efficiency	The automatic answer-saving feature makes it	32	**94.12%**

Aspect	Statement	Positive Response (f)	Percentage
	easier for me during the test.		
Learning Efficiency	The countdown timer and question navigation help me manage my test time effectively.	34	**100.00%**
Media Attractiveness	The AI-feedback website is attractive and makes me more interested in learning.	31	**91.18%**
Media Attractiveness	The combination of colors, text, and layout is comfortable to view.	32	**94.12%**
Media Interactivity	The website responds quickly when I select an answer or move between questions.	32	**94.12%**
AI-Feedback Benefits	The AI-feedback helps me identify conceptual errors in my answers.	32	**94.12%**
AI-Feedback Benefits	The explanations provided by the AI are easy to understand and help me review incorrect material.	31	**91.18%**
AI-Feedback Benefits	The solution steps provided by the AI help improve my understanding.	27	**79.41%**
AI-Feedback Benefits	The misconception identification provided by the AI helps me recognize weaknesses in my understanding.	29	**85.29%**
Learning Motivation	Using the AI-feedback website makes me more motivated and active in learning.	30	**88.24%**

The results indicate that students generally responded positively to the AI-feedback website. All 15 statements obtained positive-response percentages above 75%, and therefore all statements were categorized as very good based on the criteria used in this study.

The highest responses were observed for ease of access, ease of navigation, and learning efficiency, with several indicators receiving a 100% positive response. This suggests that students were generally able to access the website, understand its basic functions, and complete the assessment activities without substantial difficulty. Similar patterns have been reported in previous JIPK studies of digital chemistry learning media, in which student responses were used as evidence of acceptance and usability of the developed learning media. For example, an Android-based chemistry learning medium reported a 97.71% student response, while a more recent JIPK study of a STEM-integrated digital learning medium reported an 83.81% positive response.

The responses related specifically to AI-feedback are particularly relevant to the present study. The statement that AI-feedback helped students identify conceptual errors received a positive response from 32 students (94.12%). Similarly, 31 students (91.18%) agreed that the AI explanations were understandable and helped them review material they had answered incorrectly. These findings provide preliminary evidence that students perceived the feedback as more than a simple indication of whether their answers were correct or incorrect.

However, the response was somewhat lower for the statement concerning the solution steps provided by AI, with 27 students (79.41%) responding positively. This was the lowest positive-response percentage among the 15 statements. Although it remained within the "very good" category, the lower proportion suggests that the explanatory or procedural component of the feedback may require further refinement. Students may find it easier to recognize an error than to follow an AI-generated explanation of the complete conceptual or procedural solution.

The statement concerning misconception identification received a positive response from 29 students (85.29%). This is an important finding because misconception identification represents the central conceptual function of the proposed system. The response suggests that most students perceived the feedback as useful for recognizing weaknesses in their understanding. Nevertheless, the fact that five students did not provide a positive response indicates that the diagnostic-feedback relationship was not equally helpful for all users.

The response related to learning motivation was also positive, with 30 students (88.24%) indicating that the use of the AI-feedback website made them more motivated and active in learning. This finding suggests that the integration of AI-feedback into a web-based assessment environment may provide an engaging learning experience. However, motivation was measured through students' self-reported responses and should therefore be interpreted as perceived motivational support**,** rather than as an objective measure of increased motivation or learning engagement.

CONCLUSION

This study developed and implemented a limited-scale proof-of-concept of a diagnosis-informed microlearning website supported by AI-feedback for stoichiometry learning. The diagnostic assessment identified varying conceptual difficulties across six subtopics, with limiting reagents showing the highest proportion of conceptual problems (67.7%). These findings were used to guide the selection of microlearning resources and the design of targeted AI-feedback.

The limited implementation showed heterogeneous students' conceptual understanding, while students generally responded positively to the AI-feedback website. These findings provide initial evidence that diagnostic information can be operationalized to

differentiate AI-supported formative feedback according to students' identified conceptual difficulties. However, the study does not establish that the approach eliminates or significantly reduces misconceptions. Further research with larger samples and pretest–posttest or controlled designs is needed to examine its effectiveness more rigorously.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the Research and Community Service Institute (LPPM), Universitas Negeri Medan, for the financial support provided through the 2026 Beginner Lecturer Research Program under Contract No. 044/UN33.8/PPKM/PDP/2026. The authors sincerely appreciate this support, which contributed to the successful implementation of the research.

REFERENCE

Aini, R. G., Ibnu, S., and Budiasih, E. (2016), Identifikasi miskonsepsi dalam materi stoikiometri pada siswa kelas X di SMAN 1 Malang melalui soal diagnostik three-tier, *J-PEK (Jurnal Pembelajaran Kimia)*, 1(2): 50–56. https://doi.org/10.17977/um026v1i22016p50-56.

Alfiyanti, R., and Sukarmin, S. (2021), The effectiveness of software as learning media to detect and reduce

misconception in stoichiometry material, *International Journal of Chemistry Education Research*, 4(2). https://doi.org/10.20885/ijcer.vol4.iss2.art3.

Alias, N. F., and Abdul Razak, R. (2023), Exploring the pedagogical aspects of microlearning in educational settings: A systematic literature review, *Malaysian Journal of Learning and Instruction*, 20(2): 267–294. https://doi.org/10.32890/mjli2023.20.2.3.

Arief, M. D., Auliah, A., and Hardin. (2021), Pengembangan E-Magazine reaksi reduksi dan oksidasi sebagai media pembelajaran kimia kelas X SMA/MA, *Jurnal Inovasi Pembelajaran Kimia*, 3(2): 148–163. https://doi.org/10.24114/jipk.v3i2.28111.

Balasundaram, S., Mathew, J., and Nair, S. (2024), Microlearning and learning performance in higher education: A post-test control group study, *Journal of Learning for Development*, 11(1): 1–14.

Bauer, E., Richters, C., Pickal, A. J., Klippert, M., Sailer, M., and Stadler, M. (2025), Effects of AI-generated adaptive feedback on statistical skills and interest in statistics: A field experiment in higher education, *British Journal of Educational Technology*, 56(5): 1735–1757. https://doi.org/10.1111/bjet.13609.

Cai, Z., Gui, Y., Mao, P., Wang, Z., Hao, X., Fan, X., and Tai, R. H. (2023), The effect of feedback on academic achievement in technology-rich learning environments (TREs): A meta-analytic review, *Educational Research Review*, 39: 100521. https://doi.org/10.1016/j.edurev.2023.100521.

Colliot, T., Krichen, O., Girard, N., Anquetil, É., and Jamet, É. (2024), What makes tablet-based learning effective? A study of the role of real-time adaptive feedback, *British Journal of Educational Technology*, 55(5): 2278–2295. https://doi.org/10.1111/bjet.13439.

Dai, W., Tsai, Y.-S., Lin, J., Aldino, A., Jin, H., Li, T., Gašević, D., and Chen, G. (2024), Assessing the proficiency of large language models in automatic feedback generation: An evaluation study, *Computers and Education: Artificial Intelligence*, 7: 100299. https://doi.org/10.1016/j.caeai.2024.100299.

Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., and De Weerdt, J. (2021), A review of automated feedback systems for learners: Classification framework, challenges and opportunities, *Computers & Education*, 162: 104094. https://doi.org/10.1016/j.compedu.2020.104094.

Eitemüller, C., Trauten, F., Striewe, M., and Walpuski, M. (2023), Digitalization of multistep chemistry exercises with automated formative feedback, *Journal of Science Education and Technology*, 32: 453–467. https://doi.org/10.1007/s10956-023-10043-2.

Ekawisudawati, E., Wijaya, M., and Danial, M. (2021), Analisis miskonsepsi peserta didik pada materi asam basa menggunakan instrumen three-tier diagnostic test, *Chemistry Education Review*, 5(1): 62–72. https://doi.org/10.26858/cer.v5i1.13315.

Fitriani, A. R., Rahmawati, A., and Lathifa, U. (2022), Phenomenology of conceptual understanding ability, representation, and student algorithm on

stoichiometry materials, *Journal of Educational Chemistry*, 4(1): 9–18. https://doi.org/10.21580/jec.2022.4.1.9744.

Gani, A. (2025), Analisis miskonsepsi siswa pada materi stoikiometri menggunakan tes diagnostik three-tier, *BINTANG: Jurnal Pendidikan dan Sains*, 7(1): 95–111.

Hasniyah, F., and Muchtar, Z. (2021), Pengembangan uji instrumen tiga tingkat dengan CRI untuk mendeteksi miskonsepsi dalam pembelajaran reaksi redoks, *Jurnal Inovasi Pembelajaran Kimia*, 3(2): 123–135. https://doi.org/10.24114/jipk.v3i2.26503.

Huang, Y., Hansen, M., Richter, E., Kleickmann, T., Scheiter, K., and Richter, D. (2025), Enhancing preservice teachers' noticing via adaptive feedback in a virtual reality classroom, *Learning and Instruction*, 95: 102053. https://doi.org/10.1016/j.learninstruc.2024.102053.

Karimatunnisa', K., and Aloysius, H. P. (2023), Identifikasi miskonsepsi konsep stoikiometri pada sumber belajar kimia SMA, *Jurnal Riset Pembelajaran Kimia*, 8(2). https://doi.org/10.21831/jrpk.v8i2.20632.

Kinder, A., Briese, F. J., Jacobs, M., Dern, N., Glodny, N., Jacobs, S., and Lessmann, S. (2025), Effects of adaptive feedback generated by a large language model: A case study in teacher education, *Computers and Education: Artificial Intelligence*, 8: 100349. https://doi.org/10.1016/j.caeai.2024.100349.

Krause, A., Polarz, S., Hoppe, A., Ewerth, R., and Nehring, A. (2026), Towards defining, assessing and modelling competency levels in stoichiometry, *Chemistry Education Research and Practice*, 27: 351–367. https://doi.org/10.1039/D5RP00077G.

Kurniawan, C., and Jahro, I. S. (2021), Pengembangan handout titrasi asam-basa berbasis Android terintegrasi model Discovery Learning dan soal-soal HOTS, *Jurnal Inovasi Pembelajaran Kimia*, 3(2): 136–147. https://doi.org/10.24114/jipk.v3i2.28207.

Lawshe, C. H. (1975), A quantitative approach to content validity, *Personnel Psychology*, 28(4): 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x.

Lestari, E. A., Harjito, H., Susilaningsih, E., and Wijayati, N. (2021), Analisis miskonsepsi menggunakan tes diagnosa three-tier multiple choice pada materi stoikiometri, *Jurnal Inovasi Pendidikan Kimia*, 15(2): 2824–2830. https://doi.org/10.15294/jipk.v15i2.29171.

Liao, M., Zhu, K., and Wang, G. (2024), Can human-machine feedback in a smart learning environment enhance learners' learning performance? A meta-analysis, *Frontiers in Psychology*, 14: 1288503. https://doi.org/10.3389/fpsyg.2023.1288503.

Maier, U., and Klotz, C. (2022), Personalized feedback in digital learning environments: Classification framework and literature review, *Computers and Education: Artificial Intelligence*, 3: 100080. https://doi.org/10.1016/j.caeai.2022.100080.

Mejeh, M., Sarbach, L., and Hascher, T. (2024), Effects of adaptive feedback through a digital tool: A mixed-methods study on the course of self-regulated learning, *Education and*

Information Technologies, 29: 1–43. https://doi.org/10.1007/s10639-024-12510-8.

Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., and Fleckenstein, J. (2024), Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions, *Computers and Education: Artificial Intelligence*, 6: 100199. https://doi.org/10.1016/j.caeai.2023.100199.

Monib, W. K., Qazi, A., and Apong, R. A. (2025), Microlearning beyond boundaries: A systematic review and a novel framework for improving learning outcomes, *Heliyon*, 11(2): e41413. https://doi.org/10.1016/j.heliyon.2024.e41413.

Mubarokah, F. D., Mulyani, S., and Indriyanti, N. Y. (2018), Identifying students' misconceptions of acid-base concepts using a three-tier diagnostic test: A case of Indonesia and Thailand, *Journal of Turkish Science Education*, 15(Special Issue): 51–58. https://doi.org/10.12973/tused.10256a.

Panggabean, F. T. M., and Purba, J. (2021), Pengembangan e-modul terintegrasi media berbasis Adobe Flash CS6 untuk meningkatkan kemampuan pemecahan masalah kimia mahasiswa, *Jurnal Inovasi Pembelajaran Kimia*, 3(2): 116–122. https://doi.org/10.24114/jipk.v3i2.28108.

Polit, D. F., Beck, C. T., and Owen, S. V. (2007), Is the CVI an acceptable indicator of content validity? Appraisal and recommendations, *Research in Nursing & Health*, 30(4): 459–467. https://doi.org/10.1002/nur.20199.

Pölloth, B. (2025), High school students experimenting with computational chemistry: Design-based research on and through the "Comp-Chem-Lab", *Journal of Chemical Education*, 102(4): 1367–1379. https://doi.org/10.1021/acs.jchemed.4c01136.

Rilanty, N., and Juwitaningsih, T. (2020), Pengembangan media pembelajaran berbasis website untuk meningkatkan hasil belajar siswa pada materi kesetimbangan kimia, *Jurnal Inovasi Pembelajaran Kimia*, 2(1): 36–40. https://doi.org/10.24114/jipk.v2i1.17844.

Rosa, V., States, N. E., Corrales, A., Nguyen, Y., and Atkinson, M. B. (2022), Relevance and equity: Should stoichiometry be the foundation of introductory chemistry courses?, *Chemistry Education Research and Practice*, 23: 662–685. https://doi.org/10.1039/D1RP00333J.

Sa'adah, N., Haryani, S., and Mahatmanti, W. (2022), Analysis of chemical misconceptions of stoichiometry materials using Certainty of Response Index, *International Journal of Active Learning*, 7(1): 94–102.

Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., and Fischer, F. (2023), Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning, *Learning and Instruction*, 83: 101620. https://doi.org/10.1016/j.learninstruc.2022.101620.

Shiddiqi, M. H. A., Arthamena, V. D., Ayyubi, M., Manarisip, A. J., and Aznam, N. (2024), Systematic literature review: Analysis of

misconception problems and diagnostic instruments for learning chemistry, *Jurnal Penelitian Pendidikan IPA*, 10(4): 168–179. https://doi.org/10.29303/jppipa.v10i4.5189.

Silaban, R., and Panggabean, M. V. (2022), Pengembangan media pembelajaran berbasis Android pada materi kesetimbangan kimia, *Jurnal Inovasi Pembelajaran Kimia*, 4(1): 1–9. https://doi.org/10.24114/jipk.v4i1.24085.

Silaban, R., and Sianturi, P. A. (2021), Pengembangan media pembelajaran berbasis Android pada materi laju reaksi, *Jurnal Inovasi Pembelajaran Kimia*, 3(2): 191–200. https://doi.org/10.24114/jipk.v3i2.22814.

Soeharto, S., Csapó, B., Sarimanah, E., Dewi, F. I., and Sabri, T. (2019), A review of students' common misconceptions in science and their diagnostic assessment tools, *Jurnal Pendidikan IPA Indonesia*, 8(2): 247–266. https://doi.org/10.15294/jpii.v8i2.18649.

Stott, A. E. (2023), Epistemological lessons from inconsistencies in teachers' errors related to use of the mole ratio in stoichiometry calculations: A cue for professional development, *Journal of Chemical Education*, 100(7): 2548–2557. https://doi.org/10.1021/acs.jchemed.2c01252.

Tan, X., Cheng, G., and Ling, M. H. A. (2025), Artificial intelligence in teaching and teacher professional development: A systematic review, *Computers and Education: Artificial Intelligence*, 8: 100355. https://doi.org/10.1016/j.caeai.2024.100355.