JKPK (JURNAL KIMIA DAN PENDIDIKAN KIMIA), Vol. 11, No. 2, 2026 pp. 354-370
Chemistry Education Study Program, Universitas Sebelas Maret ISSN 2503-4146
https://jurnal.uns.ac.id/jkpk ISSN 2503-4154 (online)

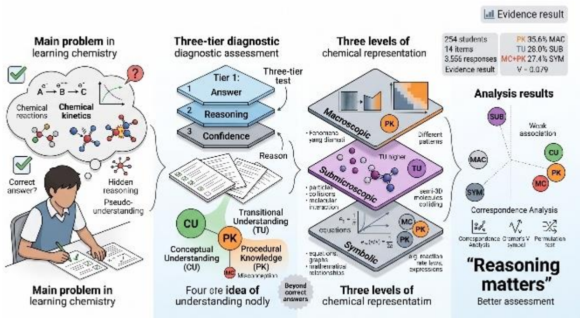
Beyond Correct Answers: Revealing Pseudo-Understanding through Correspondence Analysis of Chemical Representations in Chemical Kinetics

Suandi Sidauruk*, Ruli Meiliawati, Sisca Alvionita, Triliyansi

Department of Chemistry Education, Faculty of Teacher Training and Education, Universitas Palangka Raya, Palangka Raya, Indonesia

ABSTRACT

Correct answers in chemistry assessments do not always reflect adequate conceptual understanding, particularly when students engage with different forms of chemical representation. This issue is important because students may successfully solve problems without constructing appropriate conceptual explanations. This study investigated the distribution of Conceptual Understanding (CU), Transitional Understanding (TU), Procedural Knowledge (PK), and Misconception (MC) across macroscopic, submicroscopic, and symbolic representations in chemical kinetics. It also examined the association between representation types and understanding categories and explored aggregate patterns related to pseudo-understanding. A descriptive and associative quantitative design was employed with 254 Grade 11 students. Data were collected using 14 three-tier diagnostic test items, generating 3,556 responses. Data analysis included descriptive statistics, 10,000 restricted within-student permutation tests, Cramer's V, and Correspondence Analysis (CA). The results showed that PK was dominant in macroscopic representation (35.6%), TU was highest in submicroscopic representation (28.0%), and MC and PK occurred equally in symbolic representation (27.4% each). CU showed a relatively stable distribution across representations (20.9–23.1%). Although representation and understanding category were significantly associated ($\chi^2(6) = 44.778$, pperm < 0.001), the association was very weak (Cramer's V = 0.079; inertia = 0.013). The findings indicate that correct answers may occur without adequate conceptual reasoning and support the need for diagnostic assessments that evaluate students' reasoning and confidence alongside answer accuracy.



Keywords: *chemical representations; chemical kinetics; correspondence analysis; diagnostic assessment; pseudo-understanding; three-tier test*

*Corresponding Author: sidauruksuandi@chem.upr.ac.id

How to cite: S. Sidauruk, R. Meiliawati, S. Alvionita, and Triliyansi, "Beyond correct answers: Revealing pseudo-understanding through correspondence analysis of chemical representations in chemical kinetics," *Jurnal Kimia dan Pendidikan Kimia (JKPK)*, vol. 11, no. 2, pp. 354–370, 2026. [Online]. Available: https://doi.org/10.20961/jkpk.v11i2.118027

Received: 2026-04-26 *Accepted: 2026-08-11* *Published: 2026-08-31*

INTRODUCTION

Conceptual understanding is a central goal of chemistry education. Students need not only to produce correct answers but also to explain phenomena using relevant scientific concepts. This is challenging because chemical phenomena are understood through three levels of representation, namely macroscopic, submicroscopic, and symbolic, each of which presents information in a different form [1], [2]. Macroscopic representations describe directly observable phenomena. Submicroscopic representations explain phenomena in terms of entities and processes at the particle level, while symbolic representations express them through formulas, equations, graphs, and mathematical relationships. Meaningful understanding in chemistry therefore cannot be inferred from success with only one form of representation. Students need to connect representations accurately with the chemical concepts that underlie them [3], [4]. Differences in how learners interpret and coordinate chemical representations may influence the quality of the conceptual meaning they construct [5]. Submicroscopic and symbolic representations become particularly meaningful when learners can relate them to the chemical phenomena they are intended to explain [6].

Connecting different representations remains a challenge in chemistry learning. Students may complete tasks involving particular symbols or procedures without fully understanding the chemical meaning being represented [7]. Successful task completion should therefore be distinguished from conceptual understanding [3]. Representational competence also involves interpreting, connecting, and coordinating information presented in different forms [8], [9]. A correct answer, therefore, does not always reflect coherent conceptual understanding. This issue is particularly relevant to chemical kinetics.

Understanding reaction rates requires students to connect observable phenomena with particle collisions, activation energy, effective collisions, changes in concentration, graphs, and mathematical relationships. Learning chemical kinetics requires coordination across macroscopic, particulate, and mathematical aspects [10]. Diagnostic research has also documented persistent alternative conceptions when students reason about reaction kinetics [11]. Further evidence indicates that misconceptions in chemical kinetics may persist across different forms of diagnostic questioning [12]. Students may use equations or read graphs procedurally without being able to explain the chemical mechanisms underlying their results. Success in completing a task therefore does not necessarily reflect an equivalent level of conceptual understanding.

Diagnostic assessment allows the distinction between answer accuracy and quality of reasoning to be examined more closely. Conventional correct or incorrect scoring provides information about the accuracy of the final response, but it does not reveal the reasoning students use or how confident they are in that response. Tiered diagnostic instruments combine answers, reasoning, and confidence, providing a more detailed account of the quality of students' understanding [13]. Three-tier diagnostic tests explicitly combine content responses, reasoning, and confidence to distinguish different qualities of understanding [14]. Confidence information can also help distinguish uncertain responses from more strongly held alternative conceptions [15]. Three-tier diagnostic approaches have also been applied specifically to conceptual assessment in chemistry [16]. This approach distinguishes correct responses according to the conceptual support behind them and separates success grounded in conceptual understanding from success that is visible only in answer accuracy.

In this study, combinations of answers, reasoning, and confidence were classified into four operational categories: Conceptual Understanding (CU), Transitional Understanding (TU), Procedural Knowledge (PK), and Misconception (MC). CU represents a correct answer and correct reasoning accompanied by confidence, whereas TU represents a correct answer and correct reasoning with low confidence. PK represents a correct answer that is not supported by correct conceptual reasoning. MC includes incorrect responses that are inconsistent with scientific concepts. This classification provides more detailed information about response quality than scoring based only on correct and incorrect answers.

Distinguishing PK from CU is important when answer accuracy is used to judge students' understanding. In PK, a correct answer indicates success on the answer tier, but incorrect reasoning shows that this success is not supported by adequate conceptual reasoning. This pattern reflects the distinction between procedural success and conceptual understanding. Students may use chemical symbols or procedures correctly without demonstrating comparable understanding of the underlying concepts [3], [6]. This mismatch between successful answering and conceptual support provides the basis for examining pseudo-understanding in this study.

Pseudo-understanding is not treated as a cognitive state that can be inferred directly from a single student response. The knowledge-in-pieces perspective suggests that elements of knowledge may be activated differently across contexts, meaning that students' responses do not always reflect a fully coherent knowledge structure [17], [18]. This study, however, did not directly measure students' cognitive structures. Pseudo-understanding is therefore used in a limited sense to interpret aggregate response patterns in which successful answers are not consistently supported by correct conceptual reasoning. The term is not used as a psychological diagnosis or as a fixed cognitive category attached to individual students.

Chemical representations provide the context for examining these patterns. In this study, macroscopic (MAC), submicroscopic (SUB), and symbolic (SYM) representations refer to the dominant representational demands of each assessment item, not to students' representational preferences. This distinction is necessary because each form of representation requires a different mode of interpretation. Representational competence involves not only using a representation but also interpreting, translating, and connecting it with relevant chemical concepts [4], [19]. The distribution of CU, TU, PK, and MC across MAC, SUB, and SYM can therefore reveal variation in response quality according to the representational demands of the items.

Previous research has extensively examined misconceptions, diagnostic assessment, and representations in chemistry education. Diagnostic studies commonly focus on identifying misconceptions or grouping levels of understanding from students' response patterns [13]. Research on representations has more often examined students' ability to use, interpret, or connect macroscopic, submicroscopic, and symbolic representations [2], [9]. Further research is still needed, however, to relate conceptually distinct response categories to the representational demands of assessment items, particularly in chemical kinetics. Such analysis can show whether correct answers that lack conceptual support occur differently across MAC, SUB, and SYM.

This gap motivated the present focus on the relationship between four understanding categories and three representational contexts. The study does not seek to establish that any particular representation causes pseudo-understanding. Instead, it examines whether the distributions of CU, TU, PK, and MC vary across MAC, SUB, and SYM and estimates the strength of that association. The presence of PK across representations is then used to identify aggregate patterns of mismatch between successful answering and conceptual support, rather than to diagnose pseudo-understanding in individual students.

The analysis also recognizes that each student responded to several items. Item-level responses were therefore not treated as independent observations for statistical inference. Pearson's chi-square was used to describe the association in the contingency table, while the primary inference relied on restricted within-student permutation with students treated as exchangeability blocks. Cramer's V was used to quantify the strength of association. Correspondence Analysis (CA) was then used to explore the relative configuration of understanding categories and representations in the contingency table [20], [21]. CA was treated as a descriptive exploratory technique, and its configuration was not used to infer causal relationships or individual cognitive states.

Against this background, the study aimed to describe the distribution of Conceptual Understanding (CU), Transitional Understanding (TU), Procedural Knowledge (PK), and Misconception (MC) across macroscopic, submicroscopic, and symbolic representations in chemical kinetics. It also tested the association between understanding categories and chemical representations while accounting for repeated responses from the same students and estimated the strength of that association. Finally, Correspondence Analysis was used to explore the association pattern between understanding categories and representations and to characterize aggregate response patterns consistent with pseudo-understanding.

METHODS

1. Research Design and Participants

This study used a descriptive and associative quantitative design to examine students' understanding categories across different chemical representations in chemical kinetics. Descriptive analysis was used to characterize the distribution of understanding categories, while association analysis tested the relationship between understanding categories and the representational demands of the items. Correspondence Analysis (CA), which was also included in the earlier analytical design, was retained as an exploratory technique for mapping patterns of association in the contingency table [20], [21]. CA was not used to identify latent structures or students' cognitive states.

The participants were 254 Grade 11 students from two senior high schools in Palangka Raya, Indonesia, who had studied chemical kinetics. School A contributed 129 students and School B 125 students. The participants were distributed across eight classes, four from each school, with approximately 31 to 33 students per class. Each student was assigned a unique ID. Student, school, and class information was retained throughout data preparation and analysis to preserve the original clustering structure and identify repeated responses from the same student.

2. Diagnostic Instrument

Data were collected using a three-tier diagnostic test on chemical kinetics. The diagnostic test was designed to identify the quality of students' understanding from their response patterns rather than from final-answer accuracy alone. Diagnostic approaches in science education allow students' conceptions to be examined through the relationship between their answers and the reasoning underlying them [22], [23]. Tiered instruments extend this approach by incorporating confidence, allowing answers, reasoning, and confidence to be considered together [13]. Diagnostic instruments have also been developed to examine students' explanations of chemical reactions across multiple levels of representation [24].

Each item consisted of three tiers. Tier 1 (T1) asked a question about a chemical kinetics concept, Tier 2 (T2) provided reasoning options supporting the T1 answer, and Tier 3 (T3) measured students' confidence in their response. T1 and T2 each contained four alternatives with one correct option. T3 classified confidence as confident or not confident. This structure allowed answer accuracy to be interpreted alongside the quality of reasoning and confidence, providing a basis for distinguishing different patterns of understanding.

Table 1. Distribution of items by chemical representation

Representation	Code	Number of Items	Number of Responses
Macroscopic	MAC	2	508
Submicroscopic	SUB	4	1,016
Symbolic	SYM	8	2,032
Total		14	3,556

Fourteen items were included in the representation-based analysis. Each item was classified according to its dominant representational demand as macroscopic (MAC), submicroscopic (SUB), or symbolic (SYM). The distribution of items and corresponding response counts across the three representational categories is presented in Table 1. Representation was treated as an item characteristic rather than a student characteristic. This distinction recognizes that different representations may serve different functions and require students to coordinate information in different ways to construct understanding [4], [25].

In addition to the 14 three-tier items, the instrument included four open-ended questions designed to elicit students' written explanations of concepts related to propositions assessed in the main instrument. U1 examined students' interpretation of reaction rate and changes in concentration over time. U2 asked students to explain changes in concentration through macroscopic, submicroscopic, and symbolic representations. U3 examined reasoning about the effects of concentration on the number of particles, collisions, and reaction rate, while U4 examined students' ability to distinguish amount of substance, concentration, and reaction rate. The four questions were administered together with the main instrument and were used as supporting data to provide context for quantitative response patterns, rather than as a separate qualitative phase.

Five validators evaluated the 14 items for alignment with the indicators, conceptual accuracy, T1 construction, the relationship between T1 and T2, representational characteristics, and clarity of language. Content validity was analyzed using Aiken's V to quantify agreement among validators regarding the relevance of each item [26]. The five validators produced Aiken's V coefficients ranging from 0.842 to 0.908. A pilot test involving 60 students yielded a Cronbach's alpha of 0.745, indicating adequate internal consistency.

3. Classification of Understanding Categories

Students' responses were classified using the combination of T1 accuracy, T2 accuracy, and T3 confidence. Information from all three tiers made it possible to distinguish responses that were equally correct according to the reasoning and confidence accompanying them, consistent with the principles of tiered diagnostic assessment used in previous research [13].

Table 2. Operational classification of understanding categories

T1	T2	T3	Category	Code
Correct	Correct	Confident	*Conceptual Understanding*	CU
Correct	Correct	Not confident	*Transitional Understanding*	TU
Correct	Incorrect	Confident	*Procedural Knowledge*	PK
Correct	Incorrect	Not confident	*Procedural Knowledge*	PK
Incorrect	Incorrect	Confident	*Misconception*	MC
Incorrect	Incorrect	Not confident	*Misconception*	MC
Incorrect	Correct	Confident	*Misconception*	MC
Incorrect	Correct	Not confident	*Misconception*	MC

Four operational categories were used: Conceptual Understanding (CU), Transitional Understanding (TU), Procedural Knowledge (PK), and Misconception (MC). CU represents a correct answer and correct reasoning accompanied by confidence. TU represents a correct answer and correct reasoning with low confidence. PK represents a correct answer that is not supported by correct reasoning, regardless of confidence. MC includes incorrect responses whether the accompanying reasoning is correct or incorrect. The operational decision rules used to assign responses to the four understanding categories are summarized in Table 2. The same classification rules were applied consistently to all responses to ensure reproducibility.

PK was not equated with pseudo-understanding. PK is an observable response category determined directly from the combination of T1 and T2, whereas pseudo-understanding is used as an interpretive concept to describe an aggregate pattern in which successful answering is not consistently supported by correct conceptual reasoning. A single PK response was therefore not taken as evidence that a student had pseudo-understanding.

4. Data Preparation

The student was the primary clustering unit in the analysis. Each of the 254 students responded to 14 items, producing 3,556 item-level responses. These responses constituted repeated measurements within students and were not treated as independent observations for statistical inference. For descriptive analysis, each student's responses were summarized by MAC, SUB, and SYM, producing 762 student-representation records. Frequencies of CU, TU, PK, and MC were then aggregated into a single observed 3 × 4 contingency table. The same table was used to calculate Pearson's chi-square, Cramer's V, and CA.

The number of items differed across the three representations. Representation-specific percentages were therefore calculated using the number of responses within each representation as the denominator. Student IDs were retained in the dataset so that dependence among responses from the same student could be accounted for in the inferential procedure.

Data consistency was checked using student ID, school, class, representation code, number of items, and the total count across the four understanding categories. These checks confirmed that each student contributed 14 responses and that aggregation across MAC, SUB, and SYM matched the structure of the instrument. The final aggregation contained 3,556 responses, each assigned to one of the four categories: CU, TU, PK, or MC.

5. Analysis of Open-Ended Responses

Responses to the four open-ended questions were used as supporting data and examined against the conceptual proposition addressed by each question. Responses conveying the same conceptual meaning were grouped under a common central idea even when students used different wording, following approaches used to group open-ended responses in chemistry education research [27], [28]. The frequency and percentage of each idea were then calculated. Incorrect responses that did not express a sufficiently clear conceptual idea were recorded as unidentified central ideas. Examples reported in the manuscript were selected to represent the most frequent patterns. The open-ended data were used to clarify the quantitative findings and were not used to determine CU, TU, PK, or MC or to diagnose pseudo-understanding in individual students.

6. Statistical Analysis

The analysis followed the three research questions. For RQ1, the frequencies and percentages of CU, TU, PK, and MC were calculated separately for MAC, SUB, and SYM. Percentages within each representation were used for comparison because the numbers of MAC, SUB, and SYM items were unequal.

For RQ2, the relationship between representation and understanding category was first summarized in a 3 × 4 contingency table and described using Pearson's chi-square statistic. The asymptotic probability from Pearson's chi-square, however, was not used as the primary basis for inference. The 3,556 responses came from 254 students who each answered 14 items, so the item-level responses did not satisfy the independence assumption.

Statistical inference was therefore conducted using restricted within-student permutation. Student ID was defined as the exchangeability block. Within each student, the numbers of CU, TU, PK, and MC responses were held constant while their allocation across the 14 item positions was permuted. The representation structure was also fixed in every permutation, with two MAC, four SUB, and eight SYM positions. Responses were never exchanged across students. These restrictions preserved the dependence structure of repeated responses and followed the principle of exchangeability blocks for non-independent data [29].

A total of 10,000 restricted permutations were generated. For each permutation, the 3 × 4 contingency table was reconstructed and Pearson's chi-square statistic was calculated. The observed statistic was then compared with the empirical null distribution formed by the 10,000 permutation statistics. This approach derives the null distribution from the specified exchangeability scheme without relying on conventional asymptotic probabilities as the primary basis for inference [30], [31].

The Monte Carlo permutation probability was calculated as:

$$p_{\text{perm}} = \frac{1 + \sum_{b=1}^{B} I\left(\chi_b^2 \geq \chi_{\text{obs}}^2\right)}{B + 1}$$

where B = 10,000, χ^2obs is the Pearson chi-square statistic from the observed data, χ^2b is the Pearson chi-square statistic from permutation b, and I(·) is an indicator function equal to 1 when χ^2b ≥ χ^2obs and 0 otherwise.

Cramer's V was reported as an effect-size measure to complement the significance test because statistical evidence of association should be distinguished from the substantive magnitude of that association [32]. Cramer's V was calculated from the same observed 3 × 4 contingency table used to calculate Pearson's chi-square.

CA was then applied to the same observed 3 × 4 contingency table, ensuring consistency across Pearson's chi-square, Cramer's V, and CA. CA was used to explore the relative association structure among MAC, SUB, SYM, and the CU, TU, PK, and MC categories through row and column profiles [20], [21]. Interpretation considered total inertia, the inertia explained by each dimension, and the relative positions of points on the correspondence map. Proximity between points was treated as a descriptive feature of the CA configuration, not as a direct measure of cognitive similarity. CA was not used to infer causality, latent structure, or individual students' cognitive states.

RQ3 was addressed by integrating the distribution of understanding categories with the configuration obtained from CA. Pseudo-understanding was interpreted cautiously as an aggregate response pattern reflecting a mismatch between successful answering and conceptual support. The CA configuration and the presence of PK were therefore used to characterize aggregate-level patterns, not to diagnose pseudo-understanding in particular students.

Descriptive analyses, contingency tables, Pearson's chi-square, Cramer's V, and Correspondence Analysis were conducted using IBM SPSS Statistics version 25. The restricted within-student permutation procedure with 10,000 permutations was conducted using R version 4.6.1.

RESULTS AND DISCUSSION

1. RQ1. Distribution of Understanding Categories across Chemical Representations

The distribution of understanding categories varied across the three representational contexts. Of the 3,556 responses analyzed, 508 were associated with macroscopic representation (MAC), 1,016 with submicroscopic representation (SUB), and 2,032 with symbolic representation (SYM). Because the number of items differed across representations, comparisons were based on percentages within each representation. The frequencies and within-representation percentages of MC, TU, PK, and CU across MAC, SUB, and SYM are presented in Table 3.

Table 3. Distribution of understanding categories by chemical representation

Representation	MC, n (%)	TU, n (%)	PK, n (%)	CU, n (%)	Total
MAC	132 (26.0)	89 (17.5)	181 (35.6)	106 (20.9)	508
SUB	195 (19.2)	284 (28.0)	302 (29.7)	235 (23.1)	1,016
SYM	556 (27.4)	476 (23.4)	557 (27.4)	443 (21.8)	2,032
Total	**883**	**849**	**1,040**	**784**	**3,556**

PK had the highest proportion in MAC at 35.6%, whereas SUB showed the highest proportion of TU at 28.0%. In SYM, MC and PK occurred in equal proportions, each accounting for 27.4%. CU remained relatively stable across the three representations, at 20.9% in MAC, 23.1% in SUB, and 21.8% in SYM. The main differences across representations therefore involved the relative distributions of PK, TU, and MC rather than CU.

The substantial proportion of PK provides information that would be missed if assessment considered answer accuracy alone. In this classification, PK indicates a correct T1 answer that is not supported by correct T2 reasoning. Some responses that would appear correct under conventional scoring therefore lacked equivalent conceptual support. This pattern is consistent with the rationale for tiered diagnostic assessment, which distinguishes answer accuracy from the conceptions underlying the response [13], [22]. Taber [23] likewise emphasized that students' responses should be interpreted in relation to their underlying conceptual structure rather than only by the final answer they produce.

The open-ended responses provided a clearer view of this mismatch. In U1, for example, 65 students (25.6%) described reaction rate as the amount of substance reacting, while 58 students (22.8%) equated concentration with reaction rate. One response stated, “Reaction rate is the amount of substance that reacts.” A similar pattern appeared in graph interpretation. Fifty-four students (21.3%) associated a high reaction rate with a steep graph, while 49 students (19.3%) assumed that a higher value on the graph indicated a faster rate. These responses show that apparently appropriate terminology or representations were not always accompanied by an accurate understanding of concentration change as a function of time.

Differences in the MAC, SUB, and SYM profiles are also consistent with the nature of chemical representations, which require students to coordinate information across different levels. Representations do more than present the same concept in different forms; they may also require different interpretive processes and conceptual connections [2], [4]. Their instructional functions may likewise differ according to how learners coordinate information across representations [25]. The higher PK in MAC or higher TU in SUB therefore should not be interpreted as evidence that a particular representation produces a particular understanding category. At this stage, the data show only that the composition of responses differs descriptively according to representational demand.

The SUB pattern is noteworthy because TU reached 28.0%, the highest proportion among the three representations. Students in this category provided both correct answers and correct reasoning but reported low confidence. Submicroscopic representation requires models to explain entities and processes that cannot be observed directly. Such demands may allow students to produce correct explanations while remaining uncertain about the knowledge they are using. Because this study did not measure the cognitive processes underlying that uncertainty, however, the TU pattern in SUB is more appropriately interpreted as a response characteristic than as evidence of a specific cognitive mechanism.

Difficulties across representational levels were also evident in U2. Forty-five students (17.7%) associated a decrease in concentration with particles becoming smaller, and 43 students (16.9%) interpreted the reduction of reactants as matter disappearing. The response “Particle A becomes smaller because its concentration decreases” illustrates how a macroscopic change was not always translated appropriately at the particle level. Another difficulty appeared when students described a decrease in the number of particles as a macroscopic explanation. These patterns support previous findings that coordinating across representational levels remains problematic in understanding chemical phenomena [2], [4]. Such difficulties are also consistent with the different functions that multiple representations may serve during learning [25].

In SYM, the equal proportions of MC and PK reflect two distinct response patterns. Some students gave incorrect answers, while others produced correct answers without correct conceptual reasoning. This pattern supports the view that the ability to use chemical symbols or procedures does not necessarily imply understanding of the concepts they represent [3]. It is also consistent with representational research showing that students need to understand relationships among symbols, models, and phenomena rather than merely use representational forms procedurally [4].

Overall, all four understanding categories occurred in MAC, SUB, and SYM. No category was unique to a single representation. The descriptive differences were therefore tested further to determine whether chemical representation and understanding category remained associated after accounting for dependence among responses from the same student.

2. RQ2. Association between Chemical Representation and Understanding Category

The distributional differences identified in RQ1 were then tested for an association between chemical representation and understanding category. The observed analysis yielded χ^2obs(6) = 44.778. Because multiple responses came from the same students, significance was evaluated using 10,000 restricted within-student permutations with students retained as exchangeability blocks. This restricted permutation approach preserves the within-cluster dependence structure during randomization [29].

The permutation analysis indicated a significant association (pperm < 0.001), but the association was very weak, with Cramer's V = 0.079. Cramer's V quantifies the strength of association between categorical variables and helps distinguish statistical significance from the magnitude of the relationship [32]. Thus, although the distributions of CU, TU, PK, and MC differed across MAC, SUB, and SYM, the differences were relatively small. Chemical representation is therefore better understood as a context associated with variation in students' responses than as a factor that determines their understanding category. The main association statistics, effect-size estimate, and summary measures from the correspondence analysis are presented in Table 4.

The low Cramer's V also places the descriptive differences in RQ1 in appropriate perspective. MAC showed a relatively higher proportion of PK, SUB a higher proportion of TU, and SYM equal proportions of PK and MC. These differences, however, did not produce a strong separation between representation type and understanding category. CU, TU, PK, and MC were distributed across all three representations, with only modest shifts in their proportions.

Table 4. Association between chemical representation and understanding category

Statistic	Result
Pearson χ^2 (observed)	44.778
df	6
Number of *restricted permutations*	10,000
*p*perm	0.00010 (< 0.001)
Cramer's V	0.079
Total inertia CA	0.013
Dimension 1	68.0%
Dimension 2	32.0%

Note* The pperm value was obtained from 10,000 restricted within-student permutations, preserving the response structure within each student.

These findings suggest that representational demands are associated with variation in students' response patterns, but the relationship is weak. This is consistent with the view that chemical representations do not function as characteristics that directly determine the quality of understanding. Students must interpret the information presented, connect it with conceptual knowledge, and coordinate meaning across representational levels [2], [4]. Multiple representations may also differ in the functions they serve during meaning construction [25]. Students encountering the same representational form may therefore still display different understanding categories.

The weak association also places an important limit on interpretation. Differences among the MAC, SUB, and SYM profiles are not strong enough to claim that a particular understanding category is characteristic of a specific representation. Representation is better viewed as a context of item demands associated with response variation rather than as a determinant of students' understanding categories. This interpretation is also consistent with RQ1, where all categories appeared across all three representations.

RQ2 therefore yields two findings that need to be considered together. First, the distributions of understanding categories are not identical across MAC, SUB, and SYM, as indicated by the significant restricted permutation result. Second, Cramer's V of 0.079 shows that the strength of this association is very weak. The next question is therefore not whether an association exists, but how the relative differences among representation and category profiles are configured. This pattern was examined further through CA in RQ3.

3. RQ3. Association Patterns and Pseudo-Understanding

CA was used to clarify the relative structure of the association identified in RQ2. The total inertia of 0.013 indicates that the association structure in the contingency table is small, consistent with Cramer's V of 0.079. The two CA dimensions accounted for all of the inertia, with Dimension 1 explaining 68.0% and Dimension 2 explaining 32.0%. The two-dimensional correspondence map showing the relative configuration of chemical representations and understanding categories is presented in Figure 1.

The correspondence map shows MAC positioned relatively close to MC, SUB to TU, and SYM in a configuration relatively close to PK and MC. CU lies closer to the center of the configuration. In CA, positions near the center indicate profiles closer to the average profile and generally contribute less to the dimensions that differentiate row and column profiles [20], [21]. The position of CU therefore does not imply that students with CU responses possess integrated understanding across all representations. It indicates only that the relative distribution of CU differentiates MAC, SUB, and SYM less than some of the other categories do.

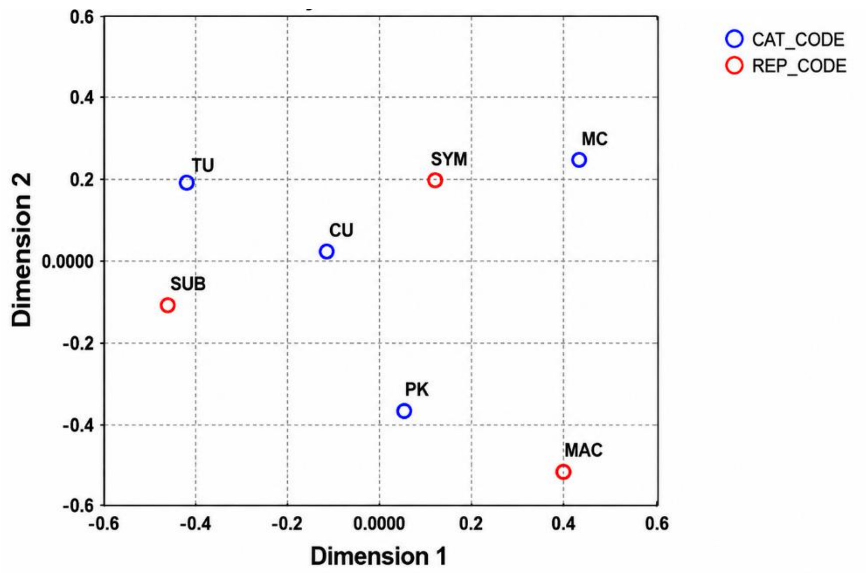


Figure 1. Correspondence analysis map of chemical representations and understanding categories.

The CA configuration should be interpreted alongside the distribution in Table 3. The proximity of SUB to TU is consistent with the higher TU proportion in SUB, at 28.0%, while the SYM configuration reflects the equal proportions of MC and PK, each at 27.4%. Proximity between points on the map does not indicate causality or similarity in cognitive processes. CA describes only the relative similarity of profiles in the contingency table [20].

The pattern most relevant to pseudo-understanding appears in PK, defined as a correct T1 answer accompanied by incorrect T2 reasoning. PK accounted for 35.6% of MAC responses, 29.7% of SUB responses, and 27.4% of SYM responses. Thus, a mismatch between successful answering and conceptual support occurred in all three representational contexts. This pattern is consistent with the view that successful completion of chemistry tasks is not always accompanied by equivalent conceptual understanding [3].

The open-ended responses provide a more concrete illustration of this mismatch. In U3, 67 students (26.4%) believed that higher concentration increases particle energy, while 61 students (24.0%) associated it with faster particle motion. One student wrote, “Higher concentration → higher energy.” This response shows that the relationship between concentration and rate may be recognized while the supporting explanation remains inconsistent with collision theory. An increase in concentration is associated with more particles per unit volume and a greater likelihood of collisions, not with an increase in particle energy.

A similar pattern appeared in U4. Sixty-four students (25.2%) believed that the amount of substance directly determines reaction rate, while 61 students (24.0%) stated that the rate remains high as long as reactants have not been depleted. The response “If there is still a lot of substance, the reaction rate must still be high” shows that amount of substance, concentration, and reaction rate were not consistently distinguished. Findings from U3 and U4 therefore show that students may recognize certain relationships in chemical kinetics while supporting them with ideas that remain conceptually inappropriate.

PK, however, cannot be equated with pseudo-understanding. PK is an operational response category determined from the combination of T1 and T2, whereas pseudo-understanding is used here to interpret a broader aggregate pattern of mismatch. This distinction matters because one or several PK responses are insufficient to infer a stable cognitive state in an individual student. Students' knowledge and reasoning may vary with the context of the task [17], [18]. Taken together, the three-tier results, CA, and open-ended responses provide complementary evidence of a mismatch between successful answering and conceptual support. Evidence relevant to pseudo-understanding comes primarily from the repeated occurrence of correct answers without adequate conceptual reasoning across different representational contexts, rather than from the proximity of points on the CA map or from any single student response. CA shows how the categories are distributed relatively, while the open-ended responses help clarify the ideas underlying some student responses. These findings reinforce the need for assessment that considers not only final answers but also students' reasoning and confidence [13], [22].

This interpretation must nevertheless be considered in light of the magnitude of the observed relationship. Cramer's V of 0.079 and total inertia of 0.013 indicate a very weak association between representation and understanding category. Representation is therefore better understood as a context in which variation in response quality appears, rather than as a factor that determines students' understanding. Within this limit, pseudo-understanding in the present study describes an aggregate pattern of mismatch between successful answering and conceptual support across representational contexts, not a fixed characteristic of students or a direct consequence of a particular representation.

CONCLUSION

This study shows that students' understanding profiles vary with the representational demands of chemical kinetics tasks. PK was more prominent in macroscopic representation, TU in submicroscopic representation, while MC and PK occurred in equal proportions in symbolic representation. CU remained relatively stable across the three representations and accounted for only about one fifth of responses. The open-ended responses further showed that some students continued to use inappropriate ideas when explaining reaction rate, changes in concentration, and particle-level phenomena. These findings indicate that correct answers are not always supported by adequate conceptual understanding.

The association between representation and understanding category remained detectable after within-student response dependence was accounted for through restricted within-student permutation, $\chi^2(6) = 44.778$, $p_{perm} < 0.001$. The association, however, was very weak (Cramer's V = 0.079; inertia = 0.013). Representation should therefore be understood as a context associated with variation in responses rather than as a determinant of a particular understanding category.

The occurrence of PK across all three representations, together with the patterns found in the open-ended responses, indicates a mismatch between successful answering and conceptual support. This pattern supports an aggregate-level interpretation of pseudo-understanding but cannot be used to diagnose individual students' cognitive states. The findings highlight the need for assessments that consider not only final answers but also students' reasoning and confidence. Future research should examine the consistency of these patterns across tasks, contexts, and time using stronger evidence at the individual level.

Author Contributions: S. Sidauruk was responsible for the formulation of the research objectives, development of the research framework, instrument preparation, data analysis, interpretation of findings, and manuscript writing. R. Meiliawati contributed to the methodological design, validation of the research instrument, review of statistical analysis, and refinement of the manuscript. S. Alvionita contributed to data collection, data processing, organization of research results, and preparation of supporting materials. Triliyansi contributed to the analysis of educational implications, literature review, and critical evaluation of the manuscript. All authors participated in the discussion of the results and approved the final version of the manuscript.

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this manuscript.

Declaration of Generative AI and AI-Assisted Technologies: During the preparation of this manuscript, the authors used AI-assisted tools to support the development of the graphical abstract and language proofreading. The authors take full responsibility for the content, accuracy, interpretation, and conclusions presented in this manuscript.

REFERENCES

[1] A. H. Johnstone, "Chemistry teaching—Science or alchemy?," *J. Chem. Educ.*, vol. 74, no. 3, pp. 262–268, 1997, doi: 10.1021/ed074p262.

[2] V. Talanquer, "Macro, submicro, and symbolic: The many faces of the chemistry 'triplet'," *Int. J. Sci. Educ.*, vol. 33, no. 2, pp. 179–195, 2011, doi: 10.1080/09500690903386435.

[3] M. M. Cooper and R. L. Stowe, "Chemistry education research—From personal empiricism to evidence, theory, and informed practice," *Chem. Rev.*, vol. 118, no. 12, pp. 6053–6087, 2018, doi: 10.1021/acs.chemrev.8b00020.

[4] V. Talanquer, "The complexity of reasoning about and with chemical representations," *JACS Au*, vol. 2, no. 12, pp. 2658–2669, 2022, doi: 10.1021/jacsau.2c00498.

[5] R. B. Kozma and J. Russell, "Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena," *J. Res. Sci. Teach.*, vol. 34, no. 9, pp. 949–968, 1997, doi: 10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U.

[6] D. F. Treagust, G. Chittleborough, and T. L. Mamiala, "The role of submicroscopic and symbolic representations in chemical explanations," *Int. J. Sci. Educ.*, vol. 25, no. 11, pp. 1353–1368, 2003, doi: 10.1080/0950069032000070306.

[7] V. Taskin and S. Bernholt, "Students' understanding of chemical formulae: A review of empirical research," *Int. J. Sci. Educ.*, vol. 36, no. 1, pp. 157–185, 2014, doi: 10.1080/09500693.2012.744492.

[8] D. F. Treagust, R. Duit, and H. E. Fischer, Eds., *Multiple Representations in Physics Education*, vol. 10. Cham, Switzerland: Springer International Publishing, 2017, doi: 10.1007/978-3-319-58914-5.

[9] M. Stieff, M. Hegarty, and G. Deslongchamps, "Identifying representational competence with multi-representational displays," *Cogn. Instr.*, vol. 29, no. 1, pp. 123–145, 2011, doi: 10.1080/07370008.2010.507318.

[10] K. Bain and M. H. Towns, "A review of research on the teaching and learning of chemical kinetics," *Chem. Educ. Res. Pract.*, vol. 17, no. 2, pp. 246–262, 2016, doi: 10.1039/C5RP00176E.

[11] Y. K. Yan and R. Subramaniam, "Using a multi-tier diagnostic test to explore the nature of students' alternative conceptions on reaction kinetics," *Chem. Educ. Res. Pract.*, vol. 19, no. 1, pp. 213–226, 2018, doi: 10.1039/C7RP00143F.

[12] Habiddin and E. M. Page, "Uncovering students' genuine misconceptions: Evidence to inform the teaching of chemical kinetics," *Acta Chim. Slov.*, vol. 70, no. 2, pp. 184–195, 2023, doi: 10.17344/acsi.2022.7880.

[13] D. K. Gurel, A. Eryilmaz, and L. C. McDermott, "A review and comparison of diagnostic instruments to identify students' misconceptions in science," *EURASIA J. Math. Sci. Technol. Educ.*, vol. 11, no. 5, pp. 989–1008, 2015, doi: 10.12973/eurasia.2015.1369a.

[14] I. Caleon and R. Subramaniam, "Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves," *Int. J. Sci. Educ.*, vol. 32, no. 7, pp. 939–961, 2010, doi: 10.1080/09500690902890130.

[15] I. S. Caleon and R. Subramaniam, "Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions," *Res. Sci. Educ.*, vol. 40, no. 3, pp. 313–337, 2010, doi: 10.1007/s11165-009-9122-4.

[16] A. Cetin-Dindar and O. Geban, "Development of a three-tier test to assess high school students' understanding of acids and bases," *Procedia Soc. Behav. Sci.*, vol. 15, pp. 600–604, 2011, doi: 10.1016/j.sbspro.2011.03.147.

[17] A. A. diSessa, "A history of conceptual change research: Threads and fault lines," in *The Cambridge Handbook of the Learning Sciences*, 2nd ed., R. K. Sawyer, Ed. Cambridge, U.K.:

Cambridge University Press, 2014, pp. 88–108, doi: 10.1017/CBO9781139519526.007.

[18] S. Vosniadou, Ed., *International Handbook of Research on Conceptual Change*, 2nd ed. New York, NY, USA: Routledge, 2013, doi: 10.4324/9780203154472.

[19] T. Phajan, K. Cojorn, and T. N. N. Dien, "Exploring the representational competence of eleventh-grade students on the rates of chemical reactions," *J. Pendidik. IPA Indones.*, vol. 14, no. 2, pp. 351–363, 2025, doi: 10.15294/jpii.v14i2.24509.

[20] M. Greenacre, *Correspondence Analysis in Practice*, 3rd ed. Boca Raton, FL, USA: CRC Press, 2017, doi: 10.1201/9781315369983.

[21] E. J. Beh and R. Lombardo, "Correspondence analysis using the Cressie–Read family of divergence statistics," *Int. Stat. Rev.*, vol. 92, no. 1, pp. 17–42, 2024, doi: 10.1111/insr.12541.

[22] D. F. Treagust, "Development and use of diagnostic tests to evaluate students' misconceptions in science," *Int. J. Sci. Educ.*, vol. 10, no. 2, pp. 159–169, 1988, doi: 10.1080/0950069880100204.

[23] K. S. Taber, *Chemical Misconceptions: Prevention, Diagnosis and Cure: Theoretical Background*, vol. 1. London, U.K.: Royal Society of Chemistry, 2002.

[24] A. L. Chandrasegaran, D. F. Treagust, and M. Mocerino, "The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation," *Chem. Educ. Res. Pract.*, vol. 8, no. 3, pp. 293–307, 2007, doi: 10.1039/B7RP90006F.

[25] S. Ainsworth, "DeFT: A conceptual framework for considering learning with multiple representations," *Learn. Instr.*, vol. 16, no. 3, pp. 183–198, 2006, doi: 10.1016/j.learninstruc.2006.03.001.

[26] L. R. Aiken, "Three coefficients for analyzing the reliability and validity of ratings," *Educ. Psychol. Meas.*, vol. 45, no. 1, pp. 131–142, 1985, doi: 10.1177/0013164485451012.

[27] A. Fink, J. D. Young, N. K. Vuppala, and R. F. Frey, "Mixed-methods exploration of students' written belonging explanations from general chemistry at a selective institution," *Chem. Educ. Res. Pract.*, vol. 24, no. 1, pp. 327–352, 2023, doi: 10.1039/D2RP00166G.

[28] D. K. Walker, V. P. Bustamante, H. E. Blomgren, J. D. Edwards, and R. F. Frey, "A mixed methods study exploring students' open-ended responses about course inclusivity in General Chemistry 1," *J. Chem. Educ.*, vol. 101, no. 4, pp. 1403–1415, 2024, doi: 10.1021/acs.jchemed.3c00781.

[29] A. M. Winkler, M. A. Webster, D. Vidaurre, T. E. Nichols, and S. M. Smith, "Multi-level block permutation," *NeuroImage*, vol. 123, pp. 253–268, 2015, doi: 10.1016/j.neuroimage.2015.05.092.

[30] A. M. Winkler, O. Renaud, S. M. Smith, and T. E. Nichols, "Permutation inference for canonical correlation analysis," *NeuroImage*, vol. 220, Art. no. 117065, 2020, doi: 10.1016/j.neuroimage.2020.117065.

[31] N. W. Koning, "More power by using fewer permutations," *Biometrika*, vol. 111, no. 4, pp. 1405–1412, 2024, doi: 10.1093/biomet/asae031.

[32] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1988. doi: 10.4324/9780203771587.