Ten essential delocalization learning outcomes: How well are they achieved?

Myriam S. Carle, Romeo Junior El Issa, Nicolas Pilote, and Alison B. Flynn*

Abstract

5   Delocalization (resonance) is a concept in organic chemistry that influences the chemical reactivity, structure, and physical properties of molecules. However, the concept has proven challenging for students and the related learning outcomes had previously been only vaguely defined. We recently defined ten essential learning outcomes about delocalization that a student should be able to demonstrate by the end of a two-course organic chemistry sequence. The goal of the present study

10   was to investigate to what extent the ten LOs were achieved by students, as well as the connections between the LOs. We analyzed three exam questions related to seven of the ten LOs for the degree of achievement, common errors, and scientific reasoning. We found that students sometimes struggled to identify when delocalization could occur, that some of the LOs built on one another, and that students were more successful in drawing resonance structures when explicitly asked, but less

15   successful when the requirement was implicit or embedded within a mechanism. Our analysis of student reasoning showed that the dominant modes of reasoning were aligned with the related expectations and explanations in the course. When asked to justify the contribution of resonance structures to the resonance hybrid, most answers used a descriptive mode of reasoning; when asked to explain why a given proton was more acidic than another, most answers contained relational and

20   linear causal reasoning. Implications for research and practice are discussed.

Keywords

Chemistry education research, Organic Chemistry, Resonance Theory, Assessment, Qualitative analysis, Quantitative analysis

Introduction

25   Delocalization, or resonance, is an integral part of organic chemistry, affecting the structure, properties, and reactivity of molecules. Delocalized structures are present in the vast majority of medicines, biological structures, materials, and other compounds, with sites of delocalized electrons directly affecting the molecule's structure, properties, and reactivity. Educators have previously reported that the subject of delocalization is difficult to learn and that students struggle with the

concept.[1] Empirical studies have also revealed several alternate conceptions about resonance, such as the view that resonance structures are alternating or that resonance structures exist as an equilibrium.[2,3] An intervention focused on building and exploring the representations in delocalization (i.e., hybrid and resonance structures) showed improvement in students' conceptions of the resonance structures.[4] Similarly, students taught by an instructor who focused on the meaning and limitations of resonance structures had higher conceptual understanding of resonance compared to students in a course setting that emphasized identifying/drawing only.[3] However, there had not been published expectations about what students should know or be able to do with delocalization concepts related to structure and reactivity. Moreover, a broader study has not been conducted on students' abilities with respect to essential learning outcomes. Our work has sought to identify the essential learning outcomes (LOs) for delocalization concepts and demonstrate how these learning outcomes can be used to assess to what extent students are achieving the LOs.

Learning outcomes are the knowledge, skills, and values that students demonstrate following a learning experience such as a module, section, course, program, or degree.[5,6] Learning outcomes help focus the instructor's and learners' attention to specific outcomes (*i.e.*, what is demonstrably learned) instead of only the inputs (*i.e.*, what is taught). Clearly communicated learning outcomes also allow students to track their own learning progress.

Intended LOs are what the instructor intends for the student to learn, including preparation of learning materials and the pedagogical approach. The *enacted LOs* are what actually occurs, including how the LOs are taught, practiced, and assessed. Part of the enacted components of LOs is the learning environments for the students. The *achieved LOs* is the skills and knowledge the student can accomplished.

In recent work, we proposed ten essential learning outcomes about delocalization that a student should be able to demonstrate by the end of an organic chemistry course sequence (Figure 1).[7] That work was based on interviews with faculty, textbook analyses, and analysis of the knowledge and skills required in future courses, building on proposals on skills of drawing resonance structures[8] to also include concepts of structure, properties, and reactivity.
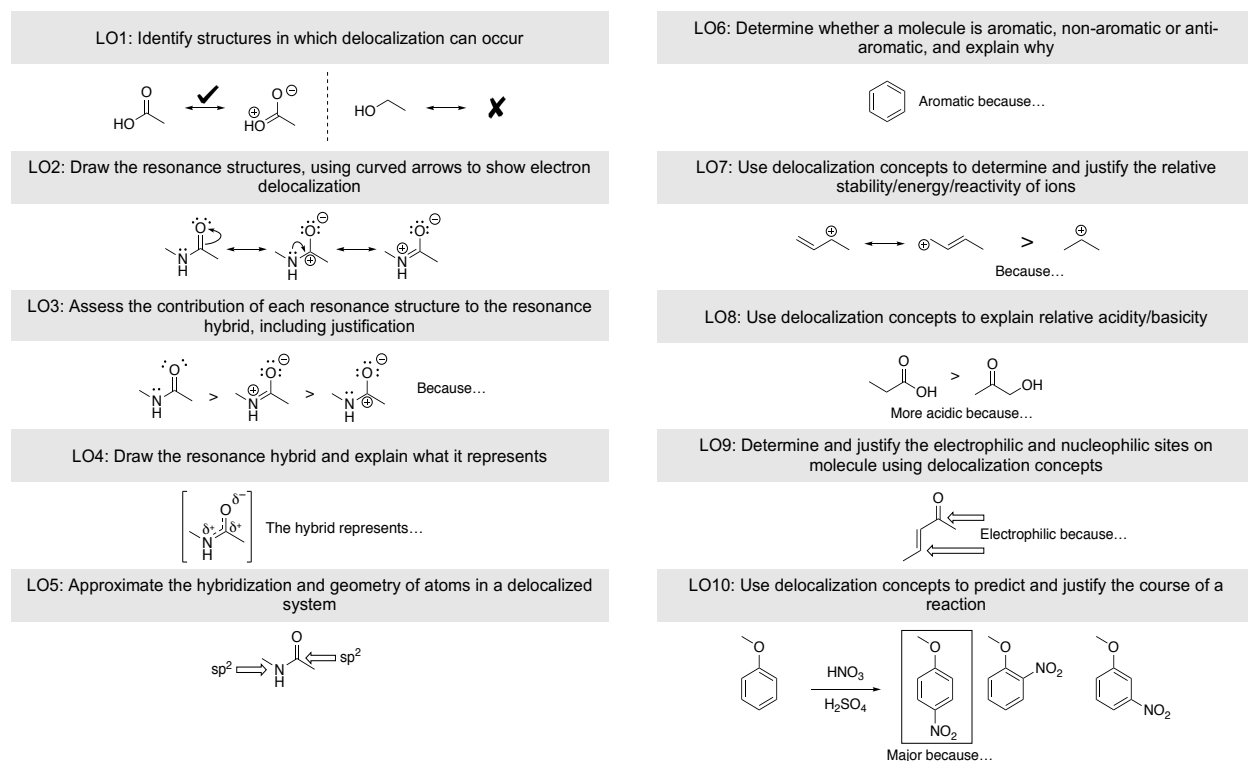
Figure 1. The ten essential LOs for delocalization. Reproduced from reference 7.

Research questions and theoretical frameworks

The goal of this research is to assess how the ten essential delocalization learning outcomes have been achieved by students in an Organic Chemistry I course at one institution.

RQ1: Which delocalization LOs are achieved more/less successfully?

RQ2: What are the common errors on delocalization-related questions?

RQ3: What is the sophistication of arguments when answering delocalization-related questions that require justification?

Theoretical frameworks

The two main theoretical frameworks used are the learning outcomes themselves (vide supra), and a reasoning framework, because several of the LOs require justifications (or arguments). A scientific argument is used to persuade and justify a claim using evidence and reasoning.[9] The *claim* is the position being argued, or the principle that is trying to be conveyed. The *evidence* is the data that is used on which the claim is based (Figure 2).[10] The *warrant* is the relationship between the claim and the evidence and why the evidence backs the claim.
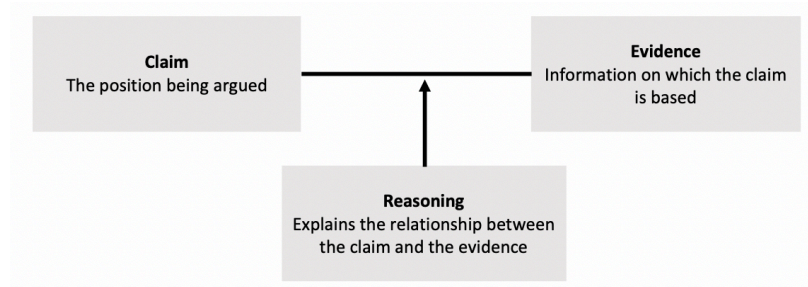
Figure 2. Toulmin's argumentation model.[10]

Several frameworks exist to qualify reasoning, such as Type I and Type II reasoning,[11,12] abstractness and abstraction,[13] rule-, case- and model-based reasoning framework,[14] levels of reasoning,[15] and mental models,[16] among others.

This research will use a framework that analyze students' responses for their level of reasoning and causality. The framework has been used and expanded in several studies.[17–20] The framework described four levels of reasoning, two non-causal (descriptive and relational) and two causal (linear and multi-component).

Descriptive reasoning occurs when entities are provided, without any further explanation. This type of reasoning is superficial and does not include causality. For example, a statement simply stating "This proton is more acidic" would be considered descriptive because the statement simply states a fact, without any relationships or why that fact is true. *Relational* reasoning occurs when a link between two entities is demonstrated; however, the underlying reason of that relationship is not explained. For example, a student stating "The proton is more acidic because of resonance" would be considered relational because the student outlined the relationship between those two concepts without explaining why those concepts are used. Causal reasoning addresses the reasons **why** a phenomenon occurs ,[13,18,21–24] and implies a cause and effect relationship between components. *Linear* causal reasoning occurs when the relation between concepts is present, and the reason is stated for why the concepts are important and how they relate to the claim. For example, consider the statement "The proton is more acidic because resonance will stabilize the conjugate base, and a more stable conjugate base means a stronger acid". In this statement, the reason is stated for why resonance is important—increased stability. *Multi-component* causal reasoning occurs when multiple linear causal relationships

are present in an argument. This type of reasoning involves weighing multiple factors and explaining why each is important, occasionally involving an analysis of why one factor is dominant.

Scientific arguments can be further analyzed to identify the granularity of an argument (e.g., global, organismal, reaction, and molecular levels) and the concepts addressed.[20,24] All aspects of an argument can vary depending on the context. For example, in some contexts, a descriptive argument at the global level that addresses two concepts could be suitable. In other contexts, molecular level granularity and multi-component causal reasoning with multiple concepts are needed to fully justify a claim. The requirements of a given contexts may be more obvious in professional/research settings; however, for students, communicating the expectations (e.g., the learning outcomes and requirements in the argument) are needed to establish the context.[23]

Methods

Settings and course

Participants in the study were taking Organic Chemistry I at a large, research-intensive Canadian university. Organic Chemistry I is offered in the winter semester of students' first year of studies, and Organic Chemistry II is offered in the summer and fall semesters. Both of these courses may be taken in either English or French and consist of two weekly lectures (1.5 hours each, mandatory, lecture or flipped format), and an optional tutorial session (1.5 hours, also called a recitation or discussion group). The Organic Chemistry I course has a required, associated laboratory section (3 hours biweekly) and the Organic Chemistry II course has a laboratory course that runs concurrently and is only required for some programs (3 hours weekly). The organic chemistry courses use a principles and patterns of mechanisms curriculum; in that curriculum, the electron-pushing formalism is explicitly taught before deeper concepts of reactivity are addressed.[25–27] The University of Ottawa's Research Ethics Board approved this study as a secondary use of data (H03-15-18).

Questions analyzed in this study

Three questions were selected to represent the intended LOs identified in previous work,[7] based on available exam questions that aligned with delocalization LOs (Figure 3). Question 1 explicitly asked students to draw the resonance structures, assess the contribution of those structures to the hybrid, draw the resonance hybrid, and label the hybridization of the atoms in the delocalized system (N = 284, Organic Chemistry I exam, 2009). The question explicitly addressed four delocalization LOs (LO2–*Draw*,

LO3–*Assess*, LO4–*Hybrid*, and LO5–*Hybridization*). Question 2 asked students to identify the most acidic proton in a molecule and justify their answer by comparing the two possible conjugate bases (N = 286, Organic Chemistry I, 2012). Delocalization was not explicitly mentioned in the question but was the main way to arrive at the correct answer. As such, the question implicitly addressed three LOs (LO1–*Identify*, LO2–*Draw,* and LO8–*Acid–base*). Question 3 asked students to draw the mechanism for the reaction between chlorobenzene and a nitronium ion (generated from HNO_2 and H_2SO_4), explicitly required students to draw resonance structures within the mechanism, then circle the most important resonance structure; Q3 addressed three LOs (LO2–*Draw*, LO3–*Assess*, without justification, and LO10–*Reaction*) (N = 178, Organic Chemistry I, 2012).
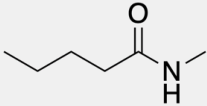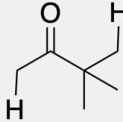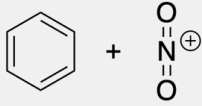
Question 1 "Resonance" (N = 280)	a. Draw all the resonance structures using the curved arrows to show electron movement. (5 points) b. Rank the resonance structures in order of contribution to the resonance hybrid and justify your answer. (6 points) c. Draw the resonance hybrid. (2 points) d. What is the hybridization of each of the following atoms? (2 points). O: _____ N: _____	
Question 2 "Acid–base" (N = 284)	a. Circle the most acidic proton in the molecule on the right. (1 point) b. Explain your answer by making a comparison of the two possible conjugate bases. (5 points)	
Question 3 "Mechanism" (N = 178)	Consider the reaction of chlorobenzene with a nitronium ion (generated from HNO_3 and H_2SO_4): a. Draw the mechanism to explain the formation of the major product, including all resonance structures. (10 points) b. Circle the most important resonance structure. (1 point)	

Figure 3. Three questions assessing 8 of the 10 LOs. Full answers can be found in the Supporting Information.

Each question was coded for level of achievement of the related LOs and common errors encountered. The full answers and coding scheme can be found in the Supporting Information (SI). The resonance question and the acid–base questions also required a justification and were therefore also coded according to level of reasoning.

Reliability of coding

The responses were coded by the first author (Questions 1 and 2) and second author (Question 3). To establish inter-rater reliability, which is the extent for which two raters agree,[28] 15% of the exams were coded by another researcher. We deemed >80% agreement and a the Krippendorf α value above 0.7 to be considered acceptable interrater reliability.[29] The percent agreement and Krippendorf α values were found to be acceptable after the first round of coding for question 2. For the other two questions agreement was not reached right away and so the raters met and discussed until agreement on the coding scheme was reached on all aspects (*i.e.*, LO achievement, errors, arguments, and reasoning). A different subset of exam was then coded and compared; the percent agreement and the calculated Krippendorf α values were deemed acceptable. All percent agreements and Krippendorf α values are listed in Table 1.

Table 1. Reliability of coding

	Question 1: Resonance		Question 2: Acid–base		Question 3: Mechanistic question	
	% Agreement	Krippendorf α	% Agreement	Krippendorf α	% Agreement	Krippendorf α
LO achievement	93%	0.85	89%	0.92	84%	0.69
Arguments	86%	0.91	96%	0.94	N/A	N/A
Reasoning	86%	0.75	81%	0.83	N/A	N/A

Results

Question 1: Draw the resonance structures

The expected answer for Question 1 (resonance) is shown in Figure 4. This answer would yield a perfect score on the exam question. Of the exams analyzed, 66% of the answers were completely correct.
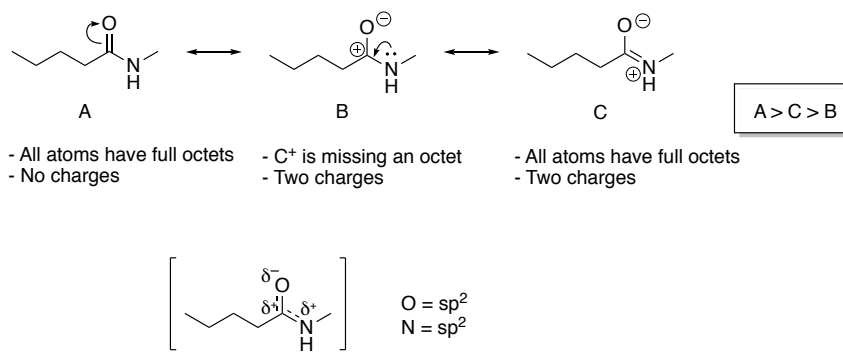

Figure 4. Expected answer for the resonance question.

LO2 (*Draw):* Most students successfully achieved this LO

LO2 (*Draw*) was demonstrated in 75% of the responses. The majority of the answers correctly contained structures **B** (90%) and **C** (86%). The types of errors found could be grouped into three categories: (1) electron-pushing formalism (EPF), (2) structures, or (3) other formalism. 15% of students had a structure error, 21% had an EPF error and 19% had another formalism error.

The majority of the errors in this question were related to the EPF. EPF errors involved incorrectly using the electron-pushing formalism (curved arrows) to demonstrate electron delocalization. These types of errors have been documented in previous work [26,27] and include reversed arrows, arrows from charges/atoms, or extra/missing arrows. The most common EPF error (80%) was related to drawing an arrow that did not start and/or finish where it should (Figure 5, left). This situation is different from a *reversed arrow*, which occurs when the arrow starts at the electron sink and ends at the source. Reversed arrows occurred in only 13% of all EPF errors.
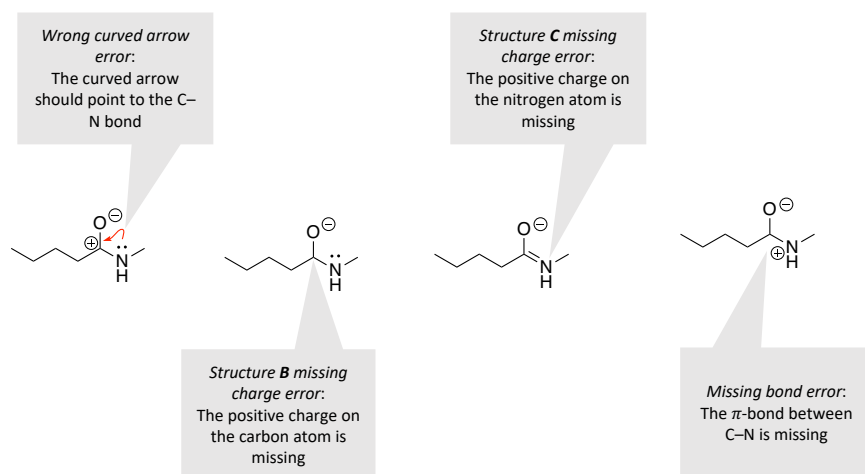

Figure 5. Most common error for each type: EPF (wrong arrow), structure B (missing charge), structure C (missing charge) and bond error (Missing bond)

Although charge- and bond-related errors occurred, they were in the minority of the exams; most answers contained the correct structures. *Structure* errors consisted of errors in drawing the resonance structures, including drawing the incorrect bonds or charges; 10% of answers had an extra structure. Most of the errors were related to drawing the charges for both structures (Figure 5, middle). The type of error differed depending on the structure. For structure **B**, the most common error was a negative charge on the carbon atom, while for structure **C**, the charge was missing on the nitrogen atom. As for bonds, the most common errors were related to the π bonds (Figure 5, right).
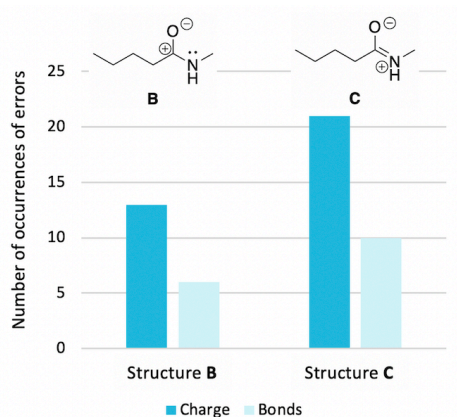
Figure 6. Question 1: Types of errors identified in each structure (N = 284).

Formalism errors are related to specific symbols to demonstrate delocalization (e.g., double headed arrow to indicate the relationship between the structures). The students' answers showed that 82% used the correct double headed arrow to show the relationship between the structures. The remaining answers included reaction arrows (i.e., one-directional), equilibrium arrows, or nothing. While this formalism is important to skillfully communicate the relationship of the structures, it was deemed not essential to achieved LO2. To demonstrate mastery of LO2, the students had to successfully draw the resonance structures and use the appropriate curved arrows.

For answers with errors, most contained only one of the error types (19%) (Figure 7), with the most common being a formalism error. The two types of errors most commonly observed together were structure and EPF errors as opposed to formalism with either of the other two, perhaps because the formalism error is not related to either of the two types. A student could be able to draw the structures and curved arrows, but not draw the appropriate formalism.
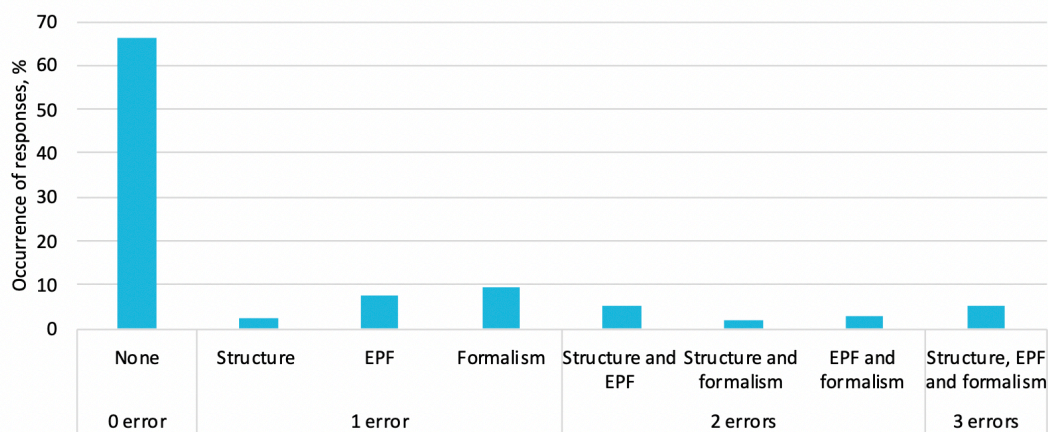
Figure 7. Occurrence of multiple errors in answers to Question 1 (N = 286).

LO3 (*Assess*): Students usually gave the correct contribution order, but often lacked the justification

LO3 (*Assess*) was fully achieved in Question 1 by 44% of the students. To fully achieve the LO, the answer had to have the correct claim (**A** > **C** > **B**) and correctly relate resonance contributor rules to back their claim, specifically: the presence of atoms with full/absent octets of electrons and number of charges.

The students' claim was correct in 68% of the exams. In the incorrect claims, 15% reversed structures C and B, but still mentioned that structure A was the major contributor. Some students (9%) had drawn an extra resonance structure, and of those only 2% had the correct claim, in which the chemically unlikely extra resonance structures were ranked last. A small minority of answers had other combinations as their claim of the relative contribution of resonance structures (3%); 5% did not contain any claim.

Six pieces of information could be leveraged as evidence to support the claim; the octet of all three structures and the charge on all three structures (Figure 8). 21% of the answers contained all six pieces of evidence. However, not all the evidence was required to justify the correct claim. For example, 51% of the students did not mention the charge on structure **B** and their reasoning was still valid. The answer could state that all atoms have a full octet in structures **A** and **C**, but that structure **B** does not have full octets, making structure B the minor contributor. Then to differentiate between structures **A** and **C**, the number of charges is used. Therefore, the evidence that structure **B** having two charges was not necessary to assign the contribution of each contributor to the resonance hybrid.
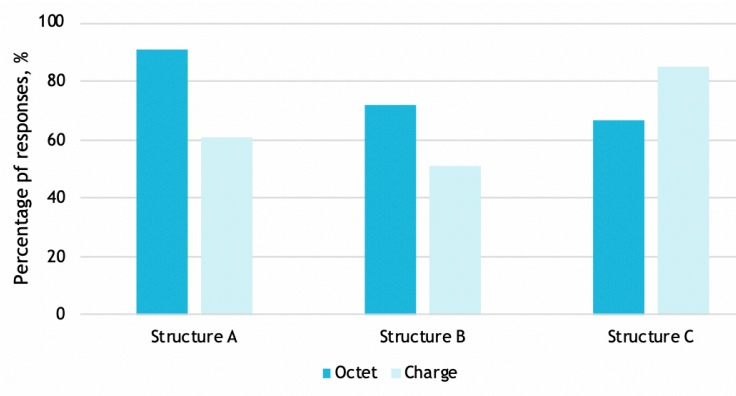
Figure 8. Evidence leveraged by students in Question 1 (resonance) (N = 280).

Most answers had a descriptive type of reasoning (Figure 9).[15,20,24] A few answers had relational reasoning, while only three answers used linear causal reasoning and none had multi-component causal reasoning. This response format matched how the topic was taught in the course.
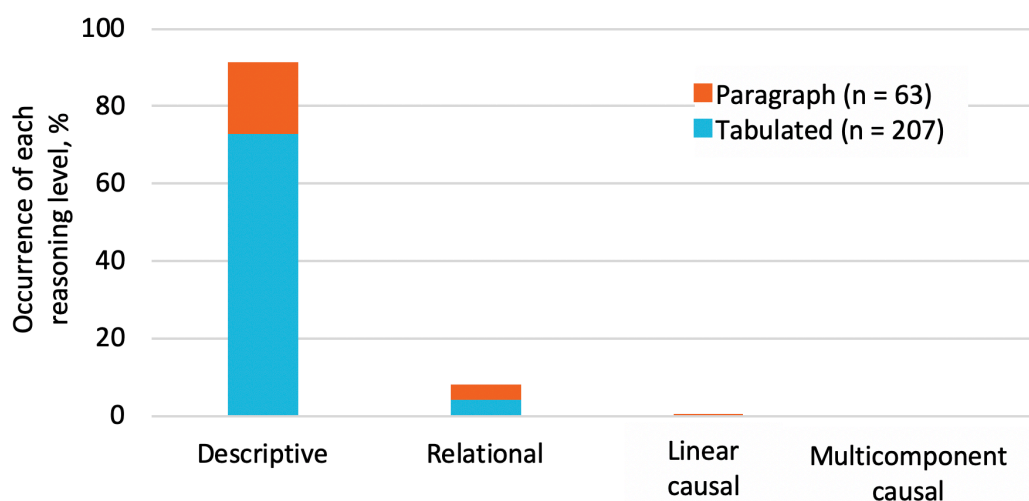

Figure 9. Levels of reasoning in Question 1 (resonance), N = 280.

The students answered the question by either listing the evidence in a tabular style or by writing out the answer in a paragraph (Figure 9). 97% of the students who wrote their answer in a tabulated (or list) format showed descriptive reasoning, as opposed to 80% of the students who used a paragraph format, representing a statistically significant difference, $X^2(df = 8, N = 280) = 300.31$, $p < 0.001$. Relational reasoning was used in 18% of the responses from paragraph answers, as opposed to 1% from table format.

LO4 (*Hybrid)*: Half of the students drew the hybrid structure correctly

Overall, 51% of the students successfully achieved LO4 and drew the resonance hybrid structure correctly. We identified three categories of errors, related to bonds, partial charges and formalism (Figure 10). Bond errors represent errors related to the dashed bonds and their location on the hybrid structure. This type error was the least common; 90% of responses had the bonds drawn correctly on the resonance hybrid structure. A partial charge error represents any mistake in the labelling the partial charges of the molecule. These types of errors could be using full charges, inversed partial charges, missing partial charges or extra partial charges. These errors were the most common (32%). The last type of error was formalism errors in which answers contained an overall charge on the resonance hybrid structure or the resonance hybrid was labelled as a transition state. Overall, most errors related to the charges and formalism.
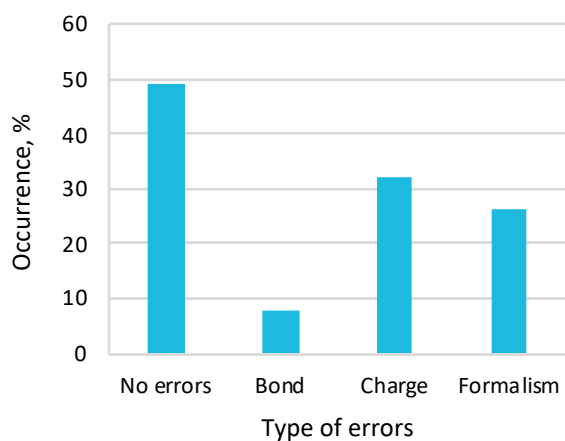

*Figure 10. Types of errors identified in the resonance hybrid structures in Question 1 (resonance). N = 280.*

Achieving LO4 (*Hybrid)* is dependent on the success of LO2 (*Draw)*. The relationship between the two LOs can be seen in Figure 11, where 97% of the students who achieved LO4 also achieve LO2. Comparatively, few students (2%) did not achieve LO2 but achieved LO4. We omitted the formalism error for both LOs because the formalism error of either would not affect the other LO. Students who successfully drew the resonance hybrid almost all drew the resonance structure correctly.

Figure 11. Connection between achievement of LO2 and LO4 in Question 1 (resonance).

LO5 (*Hybridization*): Students successfully identified the hybridization of the oxygen atom but not the nitrogen atom
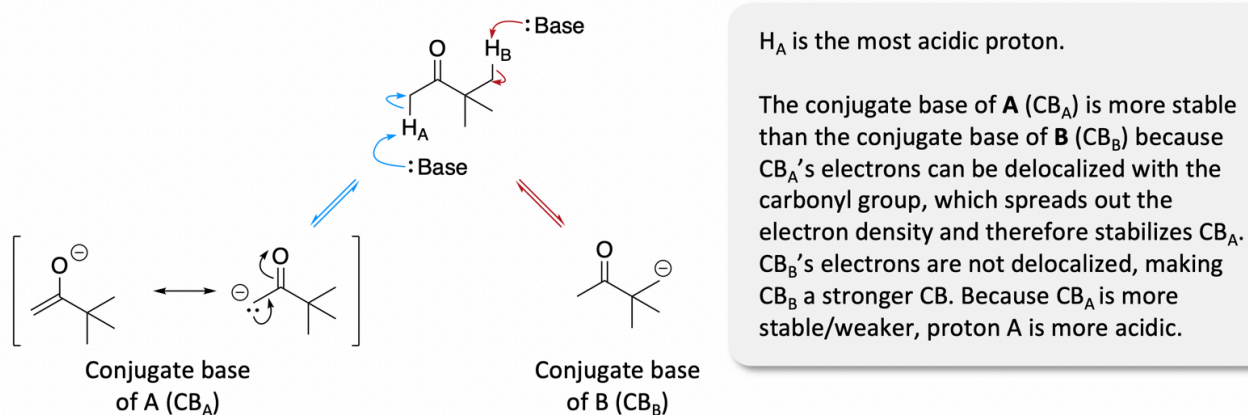
The question required students to label the hybridization of the oxygen and nitrogen atoms of an amide. Only 16% of students successfully labeled hybridization of both atoms. The oxygen was correctly identified as sp^2-hybridized by 64% of the students, while only 31% correctly identified the nitrogen amide as sp^2 hybridized (Figure 12). The nitrogen atom was most commonly labelled as sp^3-hybridized (57%). Answers of "between sp^2- and sp^3-hybridized" would be have been accepted for either atom, although no students provided that answer.
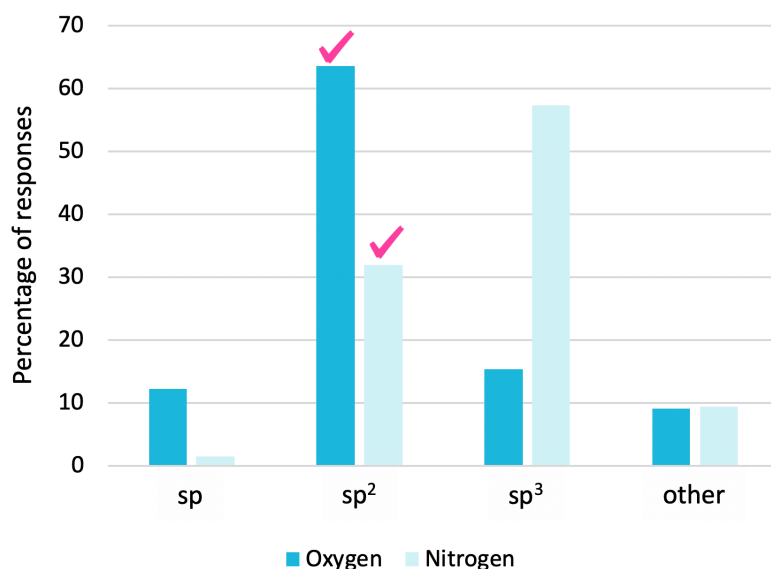
Figure 12. Hybridization labels for Question 1, part d (N = 284). The "other" category represents answers that are not represented by the three other levels, such as s or p designations.

Question 2: Acid–base question

Figure 13 shows the expected answer for Question 2, an acid–base question.

Figure 13. Expected answer for Question 2 (acid–base)

LO1 (*Identify*): Half the students identified that delocalization could occur

The question required students to identify that delocalization could occur and would stabilize the conjugate base. Conjugate base **A** could be stabilized by delocalization with and induction from the carbonyl; as such, both answers were considered acceptable. Therefore, it was not necessary for students to identify that delocalization is present in the structure to fully answer the question. For the question analyzed, 54% of answers identified that delocalization could occur.
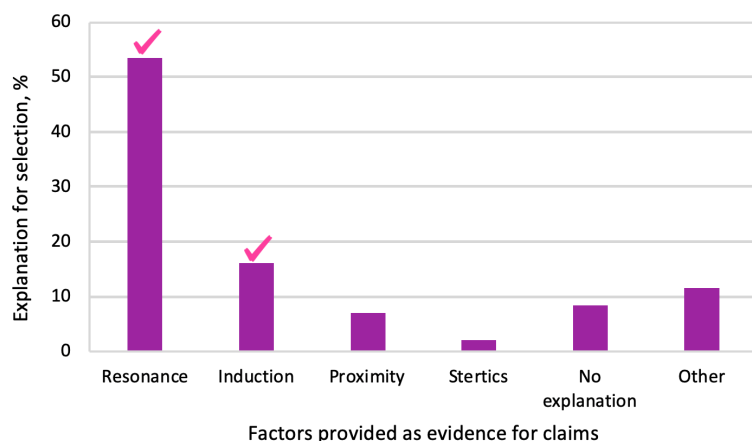

Figure 14. Factors provided as evidence for claims in Question 2 (acid–base).

LO2 (*Draw*): Errors related primarily to the wrong charge

The second question did not explicitly tell students to draw the resonance structures of the conjugate base, but it did require a comparison of the conjugate bases and resonance was possible in one of the conjugate bases. For the analysis of this LO, we only considered the answers that included resonance structures (n = 155).

Some of the students (23%) who identified that delocalization was present did not draw the resonance structures (Figure 15). These answers frequently stated that resonance was present and that it would stabilize the conjugate base, without drawing the resonance structures themselves.
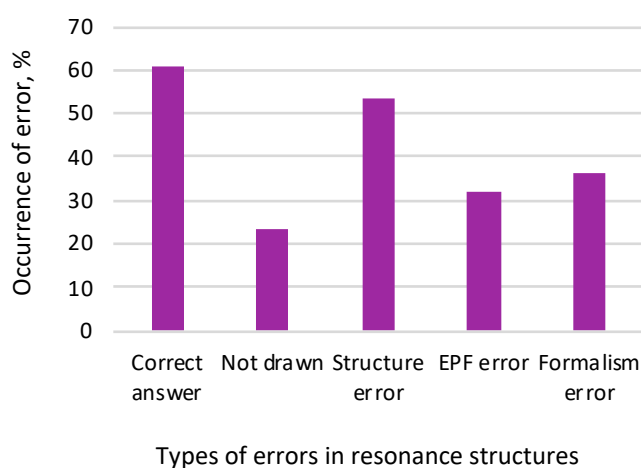

Figure 15. Type of errors in the resonance structures. (n = 155)

The most common error in the resonance structures was an incorrect charge, which is coded within the structure errors. 18% of all students drew the conjugate base with positive charges (Figure 16), despite the curved arrows of the acid-base mechanism showing electrons moving on the carbon atom. Of the students who used a resonance argument, 20% drew a positive charge.
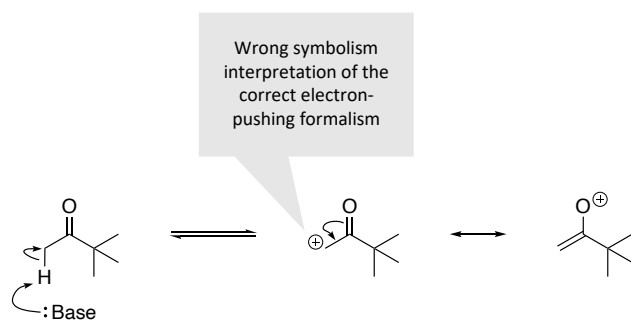

Figure 16. Question 2: Common error (18%) of drawing the wrong charge on the resonance structures, despite drawing the correct curved arrows.

For the EPF type of error, 72% of the students successfully drew the correct curved arrows to demonstrate electron delocalization, although these did not have to be drawn. The most common type of *EPF* error was a missing arrow, seen in 20% of the answers that included resonance.

The formalism type of error was minimal as 85% of students drew the correct double-headed arrow. Only 8% of the students used a reaction arrow, while 7% had no arrows between the structures.

LO8 (*Acid–base*): Less than half of the students had a causal explanation to justify their claims

We then analyzed the answers for each argument's structure, claim, evidence, and level of reasoning used. The majority of answers (72%) had the correct claim—the proton they circled in part a of the question.

The evidence used for their claim could include resonance, induction, or proximity (i.e., similar to induction, but only stating "close to oxygen"). Most of the answers that used resonance as evidence also had the correct claim (Figure 17). Other justifications, including the other correct answer of induction, had the correct claim in a lower proportion of responses.

Figure 17. Question 2 (acid–base): Evidence used to justify the claim. (N = 280)

The chemistry concepts used to justify the answers represent the evidence for the claims (Figure 18). The most five commonly-mentioned concepts were the same five concepts in the expected answer: acid strength (69%), stability of the conjugate base (64%), base strength (54%), resonance (55%), and charge (34%).
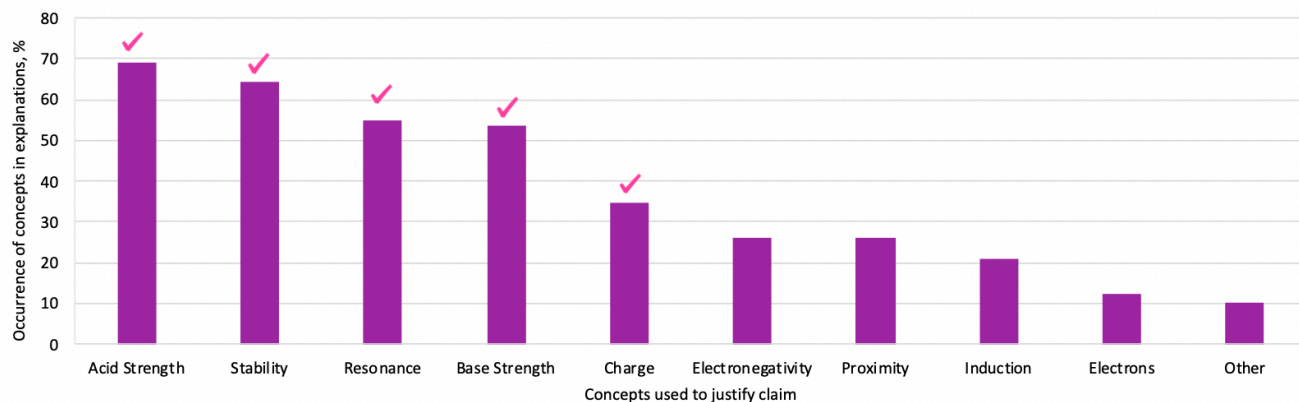
Figure 18. Question 2: Concepts used to justify claims (N = 280).

The majority of the questions showed links between concepts. Links were determined if the two concepts were connected via linking words (e.g., but, therefore, as such, for example, because) or symbols (e.g., =, <, >). Many answers (34%) had the appropriate links between all four concepts (i.e., acid strength, conjugate base strength, conjugate base stability, and resonance) (Figure 19). Many had parts of the links but lacked one of the concept or links.
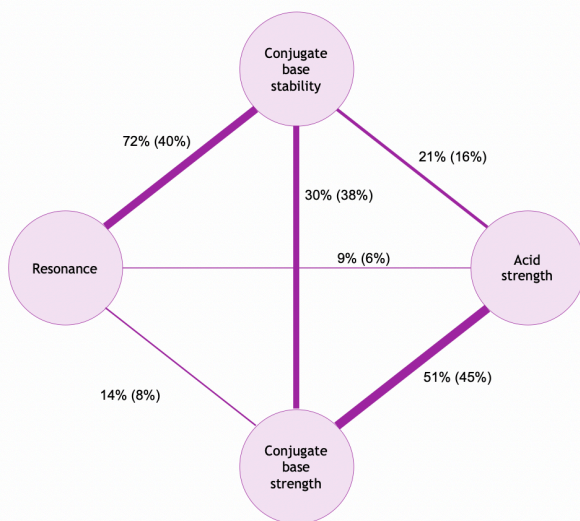
Figure 19. Percentage of answers with links between concepts (n = 155). Percentage in parentheses is out of all answers, including those who did not identify resonance (N = 286).

To be coded as causal, a response had to describe the relationship between an acid and its conjugate base (e.g., the stronger the acid, the weaker its conjugate base) **and** explain how resonance (or induction) affects the relative stability of the conjugate bases. The effect of resonance or induction

on the conjugate base needed to be shown and the argument needed to outline why that relationship
was relevant. 38% of the answers had a causal (linear-causal or multicomponent) answer, that
explained that resonance (or induction) is **why** the conjugate base is more stable (Figure 20).
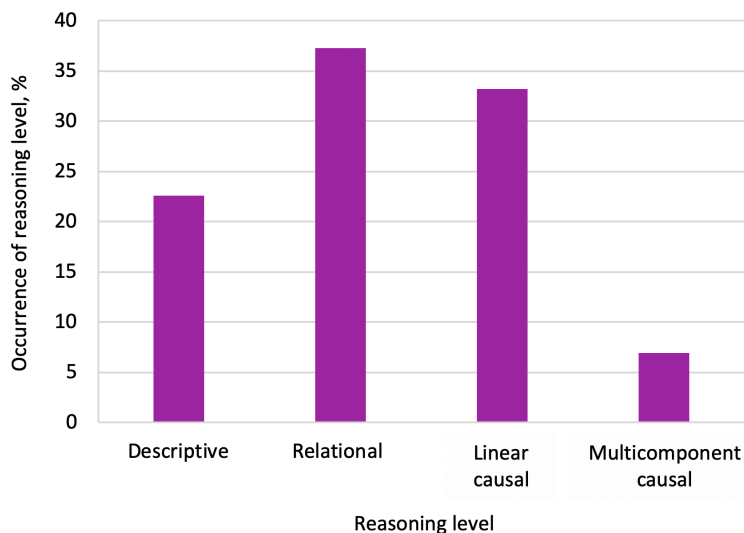

Figure 20. Level of reasoning for Question 2 (acid–base). (N = 286).

Question 3: Mechanistic question

The last question analyzed was the mechanistic question which required students to draw
resonance structure within a mechanism. The expected answer can be found in Figure 21.
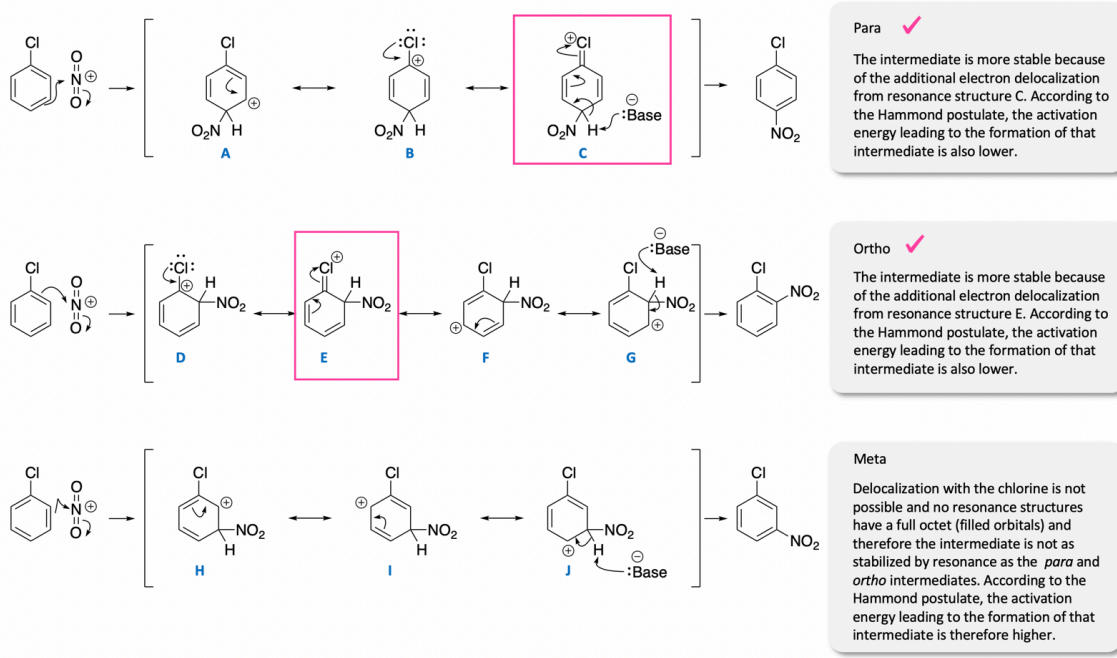
Figure 21. Expected answer for Question 3 (mechanistic).

LO2 (*Draw)*: Missing structures accounted for over 80% of the errors

LO2 (*Draw*) was achieved by only 17% of the students in this question. The question explicitly asked for resonance structures within the mechanism; however, many students (41%) did not draw any resonance structures. For the analysis of LO2, we only included answers that contained resonance structures. For the purpose of coding the achievement of this learning outcome (ability to draw resonance structures) we accepted any correctly-drawn structures. For example, if a student drew the *meta* intermediate rather than *ortho* and *para*, and drew the correct resonance structures (for *meta*), their answer was considered as having achieved LO2 (*Draw*).

The main error was a structure error; 80% of the students who drew resonance structures (N = 105) did not draw structure **C**. Other structural errors (*i.e.,* charge, bonds) were seen in less than 5% of answers that contained resonance structures. Other error types were also minimal, 20% for EPF and less than 1% for the formalism.

LO3 (*Assess)*: Students who drew structure C typically identified it as being the major resonance contributor

Few students (n = 21) drew resonance structure **C**, which makes the greatest contribution to the hybrid; of those students, 65% labelled it as the major contributor. Because the question did not require students to justify their reasoning, we could not analyze how they arrived at their conclusions.

LO10 (*Reaction*): The assessment chosen did not fully assess the LO

The results showed that 65% of all students successfully identified the *para* or *ortho* isomer as being the major contributor to the resonance hybrid. However, students may have simply memorized the rules for directing groups (i.e., that chlorine is an *ortho*/*para* director) and not used resonance in their reasoning. Most students who drew resonance structures correctly identified the *ortho* and/or *para* disubstituted benzene as being the major product (89%).

LO achievement is correlated with the final grade

We compared the number of LOs a student achieved with their final exam scores and found small to medium correlations; $r^2(282) = 0.35$ for question 1 and $r^2(175) = 0.47$ for questions 2 and 3; statistical analysis between LOs and exam scores can be found in the Supporting Information. Question 2 and 3

are used in the same correlation since they are from the same exam, while Question 1 is from a different sample of students.

Discussion

Students had varying level of achievement on the assessed LOs (Figure 22). LO1 (Identify) was
365 assessed in Question 2 (acid–base) and achieved by 54% of the students. This learning outcome is fundamental since the learning outcomes that follow often depend on being able to identify delocalization in molecules and represents a skill that is necessary in future work, such as during organic synthesis or in biochemistry. The exam question did require students to identify delocalization, but could also be answered using inductive effects in the argument. Students who did not include
370 delocalization in their answers could still have realized that delocalization could occur. A student may have been able to identify the presence of delocalization but choose not use it in their argument.
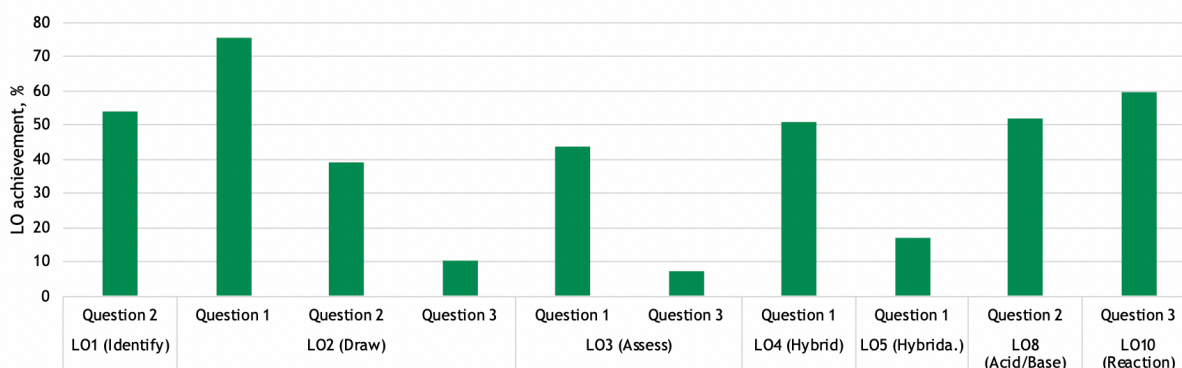

Figure 22. LO achievement across all questions. Question 1: N = 286; Question 2 (acid-base), LO1: N = 177, LO2 and LO8: n =155; Question 3 (mechanism), LO2: N = 177, LO3: n = 21, LO10: n = 105. The varying numbers are due to some answers being
375 *omitted since LOs often build on one another.*

LO2 (*Draw*) was assessed in all three questions and it was achieved to various degrees, likely depending on the type of question and the complexity of the question (Figure 22). Most students (75%) achieved the LO in Question 1 (resonance)—this question explicitly asked for resonance structures and had the lowest complexity. Fewer students (40%) achieved the LO in Question 2 (acid–
380 base)—this question did not explicitly ask for resonance structures and the question had higher complexity as resonance concepts (or inductive effects) were needed to answer a larger question. The fewest students (18%) achieved the LO on Question 3 (mechanism)—this question explicitly asked for resonance structures but in a later stage of the question and as part of a larger mechanism. We found

that students may have the skills to achieved LO2, however have difficulty using the skills when not explicitly required.

Common errors in drawing the resonance structures, were related to charges, EPF and structures which differs from previous work.[8] Previous work found that common errors were related to improbable resonance structures in which the electrons delocalized beyond the π system, breaking σ bonds, and violating an atom's octet of electrons. Within our analysis, we found this error present in few exams (9%) for Question 1. The structures analyzed were different, and so the error found in previous work could be more likely to occur in a different context. The differences could also be related how the LO was enacted by the instructor, which can lead to different outcomes.

LO3 (*Assess)* was assessed in Question 1 (resonance) and Question 3 (mechanism) but the justification was not required in the latter. For Question 2, 68% successfully achieved the LO but only 24% of the students gave a justification containing all the required evidences to justify their claims. Only 21 students achieved the LO in Question 3; the low success rate could be related to LO2 since many students did not draw the major resonance contributor. The majority of students (83%) identified structure **A** in Question 1 as the major contributor and stating all atoms had an octet and there were no charges. This differs from Betancourt-Perez and Olivera's (2010) work, which found that the most common error in selecting the most stable structure was selecting the structure for which an atom lacked an octet. However, for Question 3 of the students who drew all the resonance structures, 45% chose as the major structure one in which one atom lacked an octet of electrons. Errors in assigning the contribution of the structures toward the hybrid may be related to the structure used the assess the LO. Further instruction and assessment could address why the octet is related to stability.

Most students (88%) provided descriptive reasoning in their answer to Question 1, aligned with the expectations for that question. We also found a relationship between the mode of reasoning and the format students used to write their answers, with lower reasoning demonstrated most often by students who summarized their ideas in a table. Causality could be implied by how the table/list links the structures to the evidence; however, we decided not to make these assumptions in our coding.

LO4 (*Hybrid*) was only assessed in Question 1 (resonance) and was achieved by 64% of the students. This result varies from that of Xue and Stains (2020), who found that only 3% of students in their study

drew the resonance hybrid correctly, 14% drew a hybrid structure but missed partial charges, and more than half (53%) drew a major or minor contributor.

In the present study, the main error in drawing the hybrid related to the charges (49%), in which students either primarily drew full charges or incorrect partial charges. The locations of electron delocalization were successfully identified with dashed bonds.

The students who drew the correct resonance hybrid (LO4) had also typically achieved LO2 (*Draw*). This connection indicates that the learning outcomes are related to one another. Since a resonance hybrid is the average the resonance structures, having the correct resonance structures is important to being able to draw the hybrid. For the students who did not achieve LO2, 2% of their resonance hybrid structures were correct.

LO5 (*Hybridization*) was assessed in Question 1 (resonance) and was achieved by only 17% of the students. The students had more success identifying the hybridization of the oxygen atom (64%) in the amide than the nitrogen atom (34%). We do not know the reasoning behind those labeling since the students were not required to justify their answer. However, the results suggest that students used the structure of the first resonance structure to decide on the atoms' hybridization, rather than the more accurate hybrid structure.

LO8 (*Acid/base)* was assessed in Question 2 (acid–base) and achieved by 50% of the students*.* This learning outcome builds on LO1 and LO2, meaning that to achieve the expected answer, the previous two LOs had to be successfully achieved.

LO10 (*Reaction*) was achieved by almost 60% of the students. This LO was assessed by identifying the answers that correctly identified the regioselectivity of the reaction, which is dependent on delocalization in the first transition state, leading to the first intermediate. Since many students did not draw resonance structures, but still had the correct product, we hypothesize that students could have select the right regiochemical outcome of the reaction by simply memorizing directing group effects and drawing the correct structure. Therefore, more work would need to be done to identify the degree to which students are achieving this LO in different contexts. A scaffolded prompt that would explicitly require students to explain why a specific product is formed could potentially address the LO.

The correlation between the LOs' achievement and the students' final grades show that those who achieved multiple LOs had a higher score on the final exam grade. This relationship between LO has

been reported previously.[8] We hypothesize that the skills learned through the LOs help student conceptualize other skills (reactivity, drawing mechanisms) in organic chemistry.

When comparing answers showing success versus lack of success on a LO, the highest effect size was found with LO2 (Effect sizes: Question 1 = 1.37, Question 2 = 1.11, Question 3 = 1.98). The effect size indicates the importance of that LO on the organic chemistry exam given. LO2 requires the skill of drawing structures and using curved arrows, which is an important skill for drawing mechanisms.

Conclusions

The goal of this work was to identify the extent to which essential learning outcomes were achieved related to delocalization concepts. Achievement of the LOs varied depending on the question type, with lower success connected with more complex questions and questions in which using delocalization concepts was implicitly required.

LO1 (*Identify structures in which delocalization can occur*) was achieved by only 54% of the students on an implicit question (Q2). This result gave cause for concern, since without being able to identify delocalization the other LOs do not have the potential to be achieved. Therefore, for the other LOs to be achieved in later organic problems (e.g., organic synthesis, exploring mechanisms), being able to determine where delocalization applies is important.

The highest achieved LO was LO2 (*Draw resonance structures, using the curved arrows to show electron delocalization*), when explicitly asked as the main component of a question. However, this LO also was one of the worst achieved when it was required within a reaction mechanism. The way the LO was required seemed to affect the degree to which the LO was achieved.

We also found that some LOs built upon earlier ones. For example, students who did not identify that delocalization could occur, also could not achieve any other LO within the question. Students who achieved LO2 had a higher likelihood of achieving LO4 (*Draw the resonance hybrid*), demonstrating that the LOs are interconnected. This interconnection shows that having a strong base knowledge (LO1/LO2) is connected with higher success on the other LOs.

Student reasoning varied between the questions, which is most likely related to the question itself and to how the concepts were taught, including explanations, problem sets, and previous assessments. For Question 1 (resonance), most students used a descriptive mode of reasoning and gave their answer in a tabulated/list format, while in Question 2 (acid–base) more answers used relational and linear

causal reasoning; responses in both question types aligned with instruction in the course. Expectations and instruction could be modified in the course to include why the octet and charge are important in determining the relative contribution of resonance structures to the resonance hybrid.

Limitations

While this study analyzed students' responses on exams, we did not analyze in depth how the concepts were taught. Our investigation could not fully assess how students conceptualize and think about delocalization. This means that some student may have been able to achieve the LOs but may not have written down this information. For example, in Question 1 (resonance), students may have been able to explain their claims with causal reasoning but chose not to because of time constraints or belief of the need to. Our investigation was limited by the exams and questions available, which came from a single course section; the LOs could be assessed in other ways.

Implications for research

The link between delocalization and implicit or mechanistic questions could be investigated further. We found that students sometimes fail to recognize that delocalization occurs and had more difficulty with Questions 2 and 3, which required the connection of delocalization with other concepts. This relationship could be investigated further to determine why students did not identify that resonance was relevant, and how the link between delocalization and other concepts affects students overall organic chemistry skills. We found that some LOs are connected to being able to achieve later LOs. These links could be used to design a learning progression for the subject of delocalization, which would help align the concepts within curricula.[30]

Implications for teaching

The findings from this work could be used to inform the design and evaluation of new teaching techniques or materials. The LOs are a basis for teaching the concept of resonance that can guide instructors in their teaching and assessments. Instruction should be aligned with the desired type of reasoning (*e.g.*, relational, causal), including the taught, practiced, and assessed portions of a course or program.

Associated content

Supporting Information

Supporting Information is available.

Full answers and coding scheme (PDF)

500   Statistical analysis of LOs compared to exam scores

AUTHOR INFORMATION

Corresponding Author

Alison B. Flynn: Department of Chemistry and Biomolecular Sciences, University of Ottawa, 10 Marie

Curie Private, Ottawa, Ontario K1N 6N5, Canada

505   *E-mail: alison.flynn@uottawa.ca

Authors

Myriam S. Carle: Department of Chemistry and Biomolecular Sciences, University of Ottawa, 10 Marie

Curie Private, Ottawa, Ontario K1N 6N5, Canada

ORCID ID: https://orcid.org/0000-0003-0583-9844

510   *E-mail: mston044@uottawa.ca

Romeo El Issa: Department of Chemistry and Biomolecular Sciences, University of Ottawa, 10 Marie

Curie Private, Ottawa, Ontario K1N 6N5, Canada

Nicolas Pilote: University of Ottawa, 10 Marie Curie Private, Ottawa, Ontario K1N 6N5, Canada

Note

515   The authors declare no competing financial interest.

520   ## References:

(1) Duis, J. M. Organic Chemistry Educators' Perspectives on Fundamental Concepts and

Misconceptions: An Exploratory Study. *J. Chem. Educ.* **2011**, *88* (3), 346–350.

https://doi.org/10.1021/ed1007266.

(2) Taber, K. S. Compounding Quanta: Probing the Frontiers of Student Understanding of Molecular

525   Orbitals. *Chem. Educ. Res. Pract. Eur.* **2002**, *3* (2), 159–173. https://doi.org/10.1039/b2rp90013k.

(3) Xue, D.; Stains, M. Exploring Students' Understanding of Resonance and Its Relationship to Instruction. *J. Chem. Educ.* **2020**, *97* (4), 894–902. https://doi.org/10.1021/acs.jchemed.0c00066.

(4) Kim, T. D.; Wright, L. K.; Miller, K. An Examination of Students' Perceptions of the Kekulé Resonance Representation Using a Perceptual Learning Theory Lens. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 659–666. https://doi.org/10.1039/C9RP00009G.

(5) Collis, K. F.; Biggs, J. B. Using the SOLO Taxonomy. *Set Res. Inf. Teach.* **1986**, *2* (4), 4. https://doi.org/10.18296/set.1182.

(6) Brabrand, C.; Dahl, B. Using the SOLO Taxonomy to Analyze Competence Progression of University Science Curricula. *High. Educ.* **2009**, *58* (4), 531–549.

(7) Carle, M. S.; Flynn, A. B. Essential Learning Outcomes for Delocalization (Resonance) Concepts: How Are They Taught, Practiced, and Assessed in Organic Chemistry? *Chem. Educ. Res. Pract.* **2020**, *21* (2), 622–637. https://doi.org/10.1039/c9rp00203k.

(8) Betancourt-Ṕerez, R.; Olivera, L. J.; Rodríguez, J. E. Assessment of Organic Chemistry Students' Knowledge of Resonance-Related Structures. *J. Chem. Educ.* **2010**, *87* (5), 547–551. https://doi.org/10.1021/ed800163g.

(9) Osborne, J. F.; Patterson, A. Scientific Argument and Explanation: A Necessary Distinction? *Sci. Educ.* **2011**, *95* (4), 627–638. https://doi.org/10.1002/sce.20438.

(10) Toulmin, S. The Uses of Argument; Cambridge University Press, 1958; Vol. 34.

(11) McClary, L.; Talanquer, V. Heuristic Reasoning in Chemistry: Making Decisions about Acid Strength. *Int. J. Sci. Educ.* **2011**, *33* (10), 1433–1454. https://doi.org/10.1080/09500693.2010.528463.

(12) Talanquer, V. Concept Inventories: Predicting the Wrong Answer May Boost Performance. *J. Chem. Educ.* **2017**, *94* (12), 1805–1810. https://doi.org/10.1021/acs.jchemed.7b00427.

(13) Weinrich, M. L.; Sevian, H. Capturing Students' Abstraction While Solving Organic Reaction Mechanism Problems across a Semester. *Chem. Educ. Res. Pract.* **2017**, *18* (1), 169–190. https://doi.org/10.1039/C6RP00120C.

(14) Strickland, A. M.; Kraft, A.; Bhattacharyya, G. What Happens When Representations Fail to Represent? Graduate Students' Mental Models of Organic Chemistry Diagrams. Chem. Educ. Res. Pract. 2010, 11 (4), 293–301. https://doi.org/10.1039/C0RP90009E.

(15) Sevian, H.; Talanquer, V. Rethinking Chemistry: A Learning Progression on Chemical Thinking. *Chem. Educ. Res. Pract.* **2014**, *15* (1), 10–23. https://doi.org/10.1039/c3rp00111c.

(16) Johnson-Laird, P. N. Mental Models and Human Reasoning. *Proc. Natl. Acad. Sci.* **2010**, *107* (43), 18243 LP – 18250. https://doi.org/10.1073/pnas.1012933107.

(17) Moon, A.; Stanford, C.; Cole, R.; Towns, M. H. The Nature of Students' Chemical Reasoning Employed in Scientific Argumentation in Physical Chemistry. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 353–364. https://doi.org/10.1039/C5RP00207A.

(18) Weinrich, M.; Talanquer, V. Mapping Students' Conceptual Modes When Thinking about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2015**, *16* (3), 561–577. https://doi.org/10.1039/c5rp00024f.

(19) Moreira, P.; Marzabal, A.; Talanquer, V. Using a Mechanistic Framework to Characterise Chemistry Students' Reasoning in Written Explanations. *Chem. Educ. Res. Pract.* **2018**, *20* (2), 120–131. https://doi.org/10.1039/C8RP00159F.

(20) Bodé, N. E.; Deng, J. M.; Flynn, A. B. Getting Past the Rules and to the WHY: Causal Mechanistic Arguments When Judging the Plausibility of Organic Reaction Mechanisms. *J. Chem. Educ.* **2019**, *96* (6), 1068–1082. https://doi.org/10.1021/acs.jchemed.8b00719.

(21) Talanquer, V. Exploring Dominant Types of Explanations Built by General Chemistry Students. *Int. J. Sci. Educ.* **2010**, *32* (18), 2393–2412. https://doi.org/10.1080/09500690903369662.

(22) Cooper, M. M. Why Ask Why? *J. Chem. Educ.* **2015**, *92* (8), 1273–1279. https://doi.org/10.1021/acs.jchemed.5b00203.

(23) Bodé, N. E.; Deng, J.; Flynn, A. B. Getting past the rules and to the WHY: Causal mechanistic arguments when judging the plausibility of organic reaction mechanisms. *J. Chem. Educ.* **2019**, *96* (6), 1068–1082.

(24) Deng, J. M.; Flynn, A. B. (2020): Reasoning, Granularity, and Comparisons: A Unit-Based Method for Characterizing Students' Arguments on Chemistry Assessments. ChemRvix Prepr. 2020. https://doi.org/10.26434/chemrxiv.13119869.v1.

(25) Flynn, A. B.; Ogilvie, W. W. Mechanisms before Reactions: A Mechanistic Approach to the Organic Chemistry Curriculum Based on Patterns of Electron Flow. *J. Chem. Educ.* **2015**, *92* (5), 803–810. https://doi.org/10.1021/ed500284d.

(26) Flynn, A. B.; Featherstone, R. B. Language of Mechanisms: Exam Analysis Reveals Students' Strengths, Strategies, and Errors When Using the Electron-Pushing Formalism (Curved Arrows) in New Reactions. *Chem. Educ. Res. Pract.* **2017**, *18* (1), 64–77. https://doi.org/10.1039/C6RP00126B.

(27) Carle, M. S.; Visser, R.; Flynn, A. B. Evaluating Students' Learning Gains, Strategies, and Errors Using OrgChem101's Module: Organic Mechanisms - Mastering the Arrows. *Chem. Educ. Res. Pract.* **2020**, *21* (2), 582–596. https://doi.org/10.1039/c9rp00274j.

(28) Lange, R. T. Inter-Rater Reliability BT - Encyclopedia of Clinical Neuropsychology; Kreutzer, J. S., DeLuca, J., Caplan, B., Eds.; Springer New York: New York, NY, 2011; p 1348. https://doi.org/10.1007/978-0-387-79948-3_1203.

(29) Krippendorff, K. Reliability in Content Analysis. *Hum. Commun. Res.* **2004**, *30* (3), 411–433. https://doi.org/10.1111/j.1468-2958.2004.tb00738.x.

(30) Duncan, R. G.; Hmelo-Silver, C. E. Learning Progressions: Aligning Curriculum, Instruction, and Assessment. *J. Res. Sci. Teach.* **2009**, *46*, 606–609. https://doi.org/10.1002/tea.20316.