Investigating Instructional Strategies To Support and Elicit Organic Chemistry Students' Reasoning

by

Field M. Watts

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2023

Doctoral Committee:

Associate Professor Ginger V. Shultz, Chair
Professor Anne Ruggles Gere
Professor John Montgomery
Professor John P. Wolfe

Field M. Watts

fieldmw@umich.edu

ORCID iD: 0000-0002-1800-1816

Dedication

For Budders.

Acknowledgements

The work presented herein would not have been possible without the support and mentorship I have received from several people during my time as a graduate student. I would like to acknowledge Professor Ginger Shultz, my Ph.D. advisor and mentor, for her commitment to supporting my development as a chemistry education researcher; Professor Anne Gere, for her support and collaboration on much of the work presented herein; Professor John Wolfe and Professor John Montgomery, for their thoughtful input on my research during committee meetings; Dr. Solaire Finkenstaedt-Quinn, for being both a great friend and a great mentor—without her guidance, support, and collaboration, none of this would have been possible; and Dr. Amber Dood, for being a friend and collaborator and for supporting my growth in new areas of research.

I additionally acknowledge the researchers who I have had the opportunity to collaborate with and who have contributed to pieces of this dissertation: current and former Shultz group members Professor Jennifer Schmidt-McCormack, Ina Zaimi, and Catherine Wilhelm, along with external collaborators Professor Dr. Nicole Graulich, David Kranz, Professor Barry Thompson, Dr. Ashley Karlin, and Professor Atia Sattar. I would especially like to acknowledge the undergraduate students who I had the joy of mentoring and who contributed to this research: Michael Petterson, Paul Brandfonbrener, Grace Park, Trisha Gupte, Sabrina Archer, and Emma Snyder-White.

I would also like to thank the remaining current and former Shultz group members, for their support as thoughtful researchers, mentors, and friends: Dr. Megan Connor, Professor Eleni Zotos, Dr. Jeff Spencer, Danielle Curtis, Rebecca Fantone, Daisy Haas, Nick Garza, Dr. Blair Winograd, Professor Katy Hosbein, Cerys Rogers, Nicole Brown, Ethan Teich, Benjamin Glass, Jasen Chen, Safron Milne, and Archer Harrold. Thank you all for your valuable input and feedback on my research and for helping me think critically about my research and teaching.

Finally, I would like to thank my partner, family, and friends for their unending encouragement and support.

Table of Contents

List of Tables

List of Figures

Abstract

Success in organic chemistry requires developing reasoning skills and learning relevant concepts. Achieving these learning outcomes poses challenges for students; furthermore, supporting students with learning organic chemistry is often challenging for instructors. Because of the challenges associated with teaching and learning organic chemistry, students often engage in rote learning strategies, such as memorization, which may preclude meaningful learning. Hence, to better support students' success in organic chemistry, research is necessary to explore how novel instructional strategies can promote students' meaningful learning. Furthermore, it is important to investigate how pedagogical approaches can elicit students' reasoning—rather than simply eliciting the outcome or final solution of students' problem-solving—so instructors can better understand their students' thinking.

The research presented herein describes studies on instructional approaches intended to support students' learning and elicit students' reasoning. Most of the included research focuses on writing-to-learn, a pedagogy that uses writing assignments to support students' meaningful learning and conceptual engagement. This dissertation also provides insight into using machine learning to analyze students' responses to organic chemistry writing-to-learn assignments for the purpose of providing automated, formative feedback. This research was guided by a variety of research questions which seek to provide insights that can transform organic chemistry instruction at the undergraduate level. The dissertation opens with a review which synthesizes the existing research on students' mechanistic reasoning and how students describe and explain organic reaction mechanisms. The following two chapters address the research question of how specific instructional strategies—a mobile device application and case-comparison problems—can elicit students' reasoning about organic reaction mechanisms. The next four chapters focus on writing-to-learn, broadly addressing research questions regarding how these assignments can promote students' meaningful learning and engagement while eliciting students' reasoning about organic chemistry concepts and reaction mechanisms. The final two chapters specifically explore the question of whether machine learning models and automated feedback tools can be developed to analyze student writing and provide formative feedback for students' responses to organic chemistry writing-to-learn assignments. Several theoretical and analytical frameworks grounded the presented research, including cognitive perspectives on learning and writing, theories of engagement and meaningful learning, and mechanistic reasoning frameworks drawn from the philosophy of science literature. Interviews, surveys, and responses to writing assignments were data sources, which were analyzed using qualitative and quantitative methods.

Findings indicate how different instructional approaches can elicit different features of students' reasoning when solving problems in organic chemistry. These studies suggest that even when the outcome of students' problem-solving is the same, students often use varying approaches that may or may not be aligned with appropriate chemical thinking. Furthermore, each study reveals specific challenges students may face with different concepts and reaction mechanisms, offering implications for instructors to better support students' learning. Findings also indicate the value of writing-to-learn to promote meaningful learning in organic chemistry by supporting both cognitive and affective engagement. Results from the studies using machine learning demonstrate the feasibility of using automated analysis of students' writing to provide students with automated, formative feedback. Altogether, the studies on organic chemistry writing-to-learn assignments provide evidence of the value of this pedagogy for engaging students' learning with challenging concepts while also providing a means to elicit rich data that reflects students' reasoning. This research provides implications for instructors seeking to better elicit the depth of students' reasoning while promoting conceptual learning.

Chapter 1
Mechanistic Reasoning in Organic Chemistry

1.1 Document overview

This dissertation consists of much of my published and in-progress body of work related to investigating instructional strategies to support and elicit organic chemistry students' reasoning. The first chapter is a published review article covering the topic of mechanistic reasoning in organic chemistry, which uses a scoping review framework to encompass the organic chemistry education research literature relevant to how students describe and explain reaction mechanisms. The specific topics reviewed in the chapter include the various frameworks in the literature for characterizing students' reasoning with reaction mechanisms alongside the findings within the literature regarding how students explain features and concepts common across reaction mechanisms (specifically, the electron-pushing formalism, nucleophiles and electrophiles, acid–base theories, resonance, carbocations, and leaving groups).

The remaining chapters include my published and in-progress research articles related to how students reason with organic chemistry concepts (with a focus on reaction mechanisms), using various instructional strategies to elicit and support students' reasoning. Each chapter begins with initial remarks discussing the significance of the study, key findings related to teaching and learning, coauthor contributions, and the original publication and copyright information. The remarks also serve to provide a cohesive narrative for the body of research by discussing how the studies build upon one another. As an overview, Chapters 2 and 3 focus on two approaches to elicit and support students' reasoning with specific organic chemistry reaction mechanisms: a mobile device application to support reasoning with acid-base mechanisms and a case comparison problem structure to support reasoning with acyl transfer reactions. Chapter 4 serves as an introduction to writing-to-learn (WTL) as a pedagogical strategy which can promote students' reasoning across STEM courses; WTL is the central instructional strategy for eliciting and supporting students reasoning in the remainder of the dissertation. Specifically, the chapter reviews the research literature extending from the WTL initiative at the University of Michigan, called

MWrite. The remaining chapters focus on WTL in the context of a second-semester organic chemistry laboratory course. Chapter 5 presents a study focused on how WTL can promote students' meaningful learning experiences. Chapters 6, 7, and 8 describe studies analyzing student responses to three different WTL assignments in the organic chemistry course context, providing evidence for how WTL can promote and elicit students' reasoning about resonance and reaction mechanisms. Chapters 9 and 10 build upon the research on students' WTL responses by exploring the automated analysis of student writing using machine learning and the development of a tool to deliver automated feedback on students' drafts, respectively.

The introductory chapter was originally published as a research article in the *Journal of Chemical Education*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. Both myself and A.J. Dood are primary authors on the publication and contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing).

Original publication and copyright information

1.2 Abstract

Organic chemistry reaction mechanisms are central to the discipline of organic chemistry but challenging for students to learn and for instructors to teach. Due to the unique challenges surrounding the topic, reaction mechanisms have been the focal point of many studies within organic chemistry education research. This article provides a scoping review of the existing research on students' descriptions and explanations of organic chemistry reaction mechanisms and synthesizes the results and implications of the body of literature. The first half of the article provides an overview of the various reasoning frameworks researchers use to characterize students' reasoning with reaction mechanisms, including synthesizing the literature on students' approach to reaction mechanisms using a range of teleological, anthropomorphic, mechanistic, and causal reasoning. The second half of the article synthesizes the findings in the literature regarding students' explanations of features and concepts common across reaction mechanisms (i.e., the

electron-pushing formalism, nucleophiles and electrophiles, acid–base theories, resonance, leaving groups, and carbocations). Findings across the literature regarding approaches for supporting students' reasoning with each feature and concept are described within each section. The review concludes with a synthesis of the implications for instruction and provides an overview of directions for future research.

1.3 Introduction

Mechanisms are central to the practice of organic chemistry. Chemists use mechanisms, the electron-pushing formalism (EPF), and simultaneous application of numerous chemical and physical principles to make sense of phenomena observed in the laboratory and to describe, explain, and predict reactivity.[1,2] Introductory organic chemistry courses aim to support students' understanding of mechanisms by focusing on chemical transformations at an electronic level and supporting students' abilities to predict and explain reactivity by considering chemical properties. The epistemic importance of mechanisms in organic chemistry is underscored by the presence of mechanisms in organic chemistry textbooks and their inclusion in the undergraduate curriculum.[3–6] However, reaction mechanisms are challenging to both teach and learn, as reasoning with mechanisms requires submicroscopic understanding of molecular interactions and often requires the simultaneous integration of multiple concepts and competing properties.[7,8] The challenges with mechanisms are exacerbated by the fact that the necessary content knowledge itself (e.g., Lewis structure, acid–base theories, etc.) is often challenging for students.[6,8]

Because of the importance and associated challenges with teaching and learning reaction mechanisms, many researchers in the field of organic chemistry education research (CER) focus on eliciting students' reasoning about organic reaction mechanisms. The overarching goal of these studies is to understand how students understand and reason with reaction mechanisms. By investigating the range of students' reasoning, these studies provide insight for how instruction can better support students' movement toward expert use of mechanisms in the sensemaking process and for predicting and explaining reactions. Across these studies, students demonstrate many different types of reasoning (e.g., teleological or causal reasoning), only some of which are chemically sound. Furthermore, these studies provide details about how students understand and reason with a variety of properties (e.g., nucleophilicity and electrophilicity, acidity and basicity, etc.) and structural features (e.g., leaving groups and carbocations) relevant across reaction mechanisms. The goal of this review article is to provide an overview of the organic CER literature pertaining to how students reason when describing and explaining reaction mechanisms and when using the electron-pushing formalism. The research questions we aim to address with this review are as follows:

1. What types of reasoning do students use when describing and explaining reaction mechanisms in organic chemistry?
2. How do students describe, explain, and reason with common concepts and features across different reaction mechanisms?

1.4 Scoping review framework

This review article was conducted as a scoping review.[9–11] Scoping reviews differ from systematic reviews in that the quality of reviewed studies are not assessed; rather, the goal of a scoping review is to synthesize the range of literature related to the guiding questions of the review.[12] The methodology of a scoping review requires identifying the research question(s), iteratively identifying and selecting relevant studies, charting the findings from the selected studies, and summarizing and reporting the results.[9,10] This methodology allowed us to map the range of literature relevant to how students describe, explain, and reason with reaction mechanisms in organic chemistry. The scoping review process allows for a descriptive overview of a large body of research and to review the qualitative aspects of the relevant studies. Furthermore, the scoping review methodology is valuable for identifying gaps in the literature, suggesting directions for future research, and outlining implications for practice.

1.5 Methods

1.5.1 Positionality statement

Reviewing the broad range of literature pertaining to how organic chemistry students describe, explain, and reason with reaction mechanisms is not a straightforward task; any researcher or group of researchers may approach the task in different ways. Therefore, it is important to acknowledge our positionalities as the authors of this article. We include this positionality statement in an effort to describe our reflexivity with respect to our approach for this review article and to contribute to the trustworthiness of this work.[13] Furthermore, we provide this statement to emphasize that this review is not definitive, as synthesizing the literature requires interpretation that is influenced by our own perspectives within the field. AJD is a postdoctoral researcher at the University of Michigan with a Ph.D. in Chemistry from the University of South Florida (research focus in CER), and FMW is a Ph.D. candidate at the University of Michigan with a focus in CER and an M.S. degree in Chemistry. As early career researchers studying how students describe, explain, and reason with reaction mechanisms, we recognized the need for a synthesis of the frameworks and approaches researchers use to characterize students' reasoning with organic chemistry reaction mechanisms. Furthermore, we recognized the need for a review of the specific findings pertaining to common features and concepts across reaction mechanisms. We believe this work will be useful for other researchers, especially graduate students and researchers new to the field. It is necessary to additionally note that we are authors on 5 (AJD) and 5 (FMW) articles included within the scope of this review, and are actively pursuing research relevant to the topic of organic chemistry students' descriptions, explanations, and reasoning with reaction mechanisms. The inclusion of our own articles in this review may have enhanced our identification of where the findings within our articles align with the presented themes. To account for this potential bias, we engaged in thematic analysis through a consensus approach by which we both discussed each article included within the review throughout the analysis (as described in more detail below). Furthermore, we have made an effort to ensure that our contributions to the literature are not weighted differently from the other contributions included within the scope of the review. We additionally sought feedback from researchers with contributions in the area during the analysis and while drafting this article.

1.5.2 Scope of the included literature

This review encompasses the research literature on students' descriptions and explanations of reaction mechanisms in organic chemistry. A previous article published in 2015 synthesizes the organic chemistry education literature between 2000 and 2015, broadly focused on students' problem solving, use of the EPF, conceptual knowledge, and cognitive skills.[8] While the existing article includes a section on mechanistic problem solving, the article does not synthesize the findings throughout the literature related to how students describe and explain reaction mechanisms. Furthermore, besides describing students' conceptual knowledge of acid–base concepts, the review does not synthesize the findings pertaining to students' understanding of other concepts and features necessary for describing and explaining reaction mechanisms that are prevalent in the literature (e.g., nucleophilicity and electrophilicity, carbocations, leaving groups, etc.). As the goal of this review is to synthesize the findings pertaining to how students describe and explain reaction mechanisms, including how students' understand and incorporate key concepts and features of reaction mechanisms, this review covers the literature from the time period of the previous article (2000 through 2015). The present review also includes relevant articles in the time period since the prior article and before January 1, 2022. Additional literature not directly related to students' descriptions and explanations of mechanisms is cited where appropriate for contextualizing the synthesized findings. In this work, we did not focus on the body of literature that covers more general aspects of students' problem solving in organic chemistry.

1.5.3 Data collection

Initial articles screened for inclusion in this review were identified by searching for the key terms "organic chemistry" and "reaction mechanism" in this *Journal*, *Chemistry Education Research and Practice*, the *International Journal of Science Education*, the *Journal of Research in Science Teaching*, and the ERIC database. The titles, keywords, and abstracts of these articles were screened to identify if they contributed findings pertaining to students' descriptions and explanations of and/or reasoning with organic reaction mechanisms or directly related concepts or skills (e.g., nucleophiles and electrophiles, resonance, or representational competence in organic chemistry). Articles that were not initially clear as to whether they contributed findings pertaining to the scope of the review were subjected to further evaluation. For the potentially relevant articles, one of the authors scanned the methods, results, and discussion sections to identify whether the articles fit within the scope of the review; articles identified to fit within the scope were included in the set of articles for further evaluation. For any articles that were still ambiguous after this screening, we both more closely examined and discussed whether or not they fit within the scope of the review to reach a consensus. The set of articles identified up to this point were all subjected to further screening to identify whether they incorporated a clear research objective and/or research questions, data analysis, and results. This decision was made to limit the scope of the review to only incorporate articles with research contributions. All articles that made it through this selection process were read by both authors for inclusion in the review. During the process of reading and evaluating the selected articles, we also retrieved and screened for further inclusion any additional articles that were cited within the articles read for the review and appeared potentially relevant.

These articles were subjected to the same screening process as the articles initially retrieved and screened for inclusion. The process of initially screening articles for inclusion and then including additional articles cited within the original set of articles aligns with the objectives of the scoping review to include any literature identified as fitting within the scope of the review. In total, 73 articles were included in the review, with publication dates ranging from 2001 to 2021; Table 1.1 identifies the number of articles drawn from each journal. While we intend for this review to be thorough and comprehensive, we acknowledge the potential for unintentional omissions.

Table 1.1 Number of articles included in this review

Journal	**Number of Articles**
Chemistry Education Research and Practice	39
Journal of Chemical Education	24
International Journal of Science Education	5
Journal of Research in Science Teaching	2
Canadian Journal of Chemistry	1
International Journal of Physics and Chemistry Education	1
Learning and Instruction	1
Total	73

1.5.4 Data analysis

All articles identified for inclusion in the review were read by both authors to identify their contributions to the literature related to how students describe, explain, and/or reason with organic reaction mechanisms. Specifically, each article included was charted to note the methods used (qualitative, quantitative, or mixed methods), the study population (undergraduate students, graduate students, faculty, or practicing chemists), and the type of reaction focused on (e.g., acid–base reactions or substitution reactions). For articles which characterized students' descriptions and explanations using a framework focused on reasoning, we also noted the reasoning framework used.

We used a thematic analysis process to review the literature and address the research questions for this review article.[14] While reading the articles, we identified initial themes pertaining to the reasoning strategies students used (e.g., noting articles describing students' teleological or mechanistic reasoning) and the concepts and features included in students' descriptions and explanations of organic reaction mechanisms (e.g., noting which articles described students' understanding of nucleophiles and electrophiles or acid–base theories). During the process of discussing the articles, rereading, and charting the literature, we organized the notes into specific themes. In alignment with our research questions, the broad themes are (1) the reasoning strategies students used when describing or explaining reaction mechanisms, and (2) how students reason with concepts and features common across reaction mechanisms. The second theme is divided into subthemes focused on the prevalent concepts and features in the reviewed articles:

1. The electron-pushing formalism
2. Nucleophiles and electrophiles
3. Acid–base theories
4. Resonance
5. Carbocations
6. Leaving groups

With the list of themes pertaining to the common findings across the set of reviewed articles, we returned to each article to chart the literature by noting the findings pertaining to each theme while rereading the reviewed articles. To establish trustworthiness for the data analysis process, each article was read and charted by one author followed by the other author reading the article and checking the decisions for which themes and findings were noted.[13] For ambiguous cases or areas of disagreement with respect to whether an article included findings pertaining to a specific theme, both authors discussed the article to ensure relevant findings were noted for each theme. Through this process, both authors read every included article during the charting process and came to a consensus on which articles were charted for each theme.[13,15]

1.6 Results

The results provide a descriptive overview of the articles included in the review. Figure 1.1 describes the publication years of the reviewed articles, grouped by whether they used qualitative, quantitative, or mixed methods. The trend across the years indicates that research pertaining to how students reason about, describe, and explain organic reaction mechanisms has grown in recent years. This is an active and growing field of inquiry, indicating the value of synthesizing the findings across the set of articles in order to highlight implications for instructors and identify avenues for future research. Of note, a majority of articles used either qualitative or mixed methods. Seven of the ten articles using quantitative methodologies have appeared mostly in recent years (between 2018 and 2021); four of these focused on machine learning methodologies to analyze students' descriptions and explanations of reaction mechanisms.[16–19] Machine learning represents a new and growing area of quantitative research in the area of understanding students' reasoning with reaction mechanisms. The relatively few quantitative studies indicate that leveraging quantitative methodologies to explore how students describe, explain, and reason with reaction mechanisms is an avenue for further research.

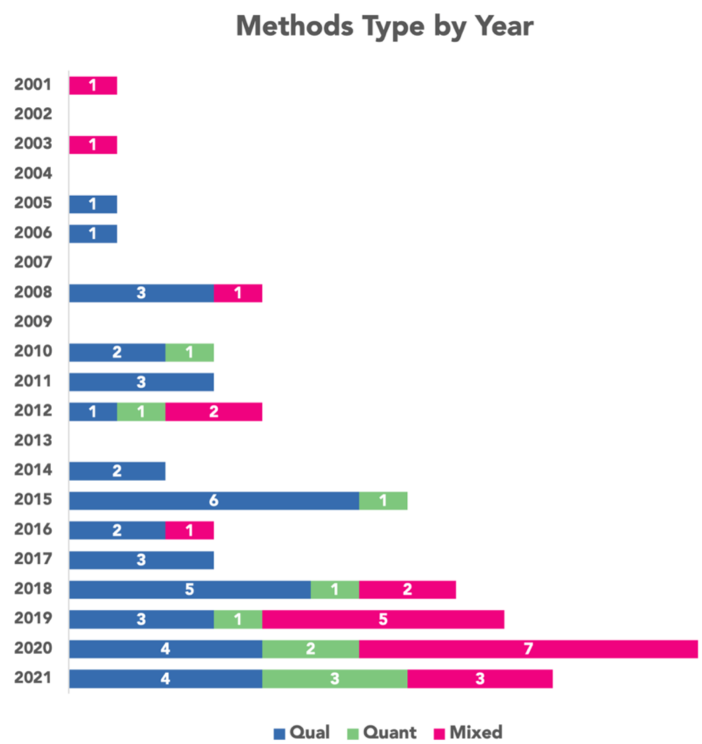


Figure 1.1 Publication years for the included articles, indicating the proportion of qualitative, quantitative, and mixed methods articles each year.

Figure 1.2 indicates the study population(s) across the articles. Of note, a majority of articles focus on undergraduate students' descriptions and explanations of reaction mechanisms, with only some articles including graduate students and even fewer including faculty and practicing chemists. This suggests that future research, in general, may seek to understand how experts and graduate students reason about, describe, and explain reaction mechanisms to more fully understand the nature of these tasks. More fully understanding expert reasoning will allow researchers and instructors to work toward more clearly defining the desired learning outcomes for undergraduate students.

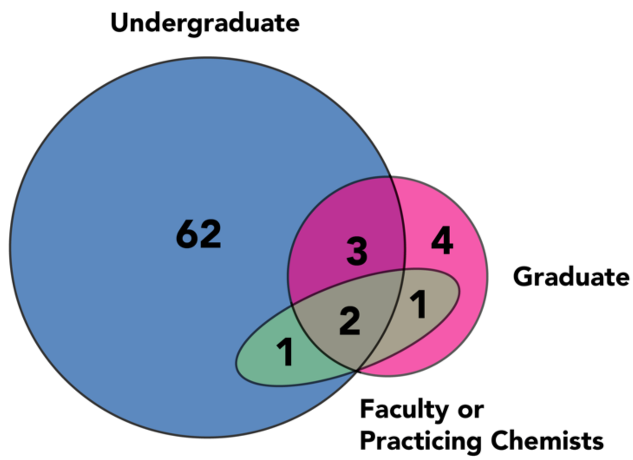


Figure 1.2 Study populations across the reviewed articles.

Figure 1.3 indicates the reaction types focused upon within the reviewed articles. A subset of articles (*n* = 16) did not focus directly on students' reasoning with a specific reaction type, and instead presented findings pertaining to specific concepts (e.g., acids and bases, nucleophiles and electrophiles, resonance) or abilities (e.g., representational necessary for mechanistic reasoning). Lastly, Figure 1.4 illustrates the proportion of reviewed articles that include information relevant to one of the themes that emerged during the thematic analysis and how the themes overlap across articles. The Discussion section of this article will review the literature pertaining to each of the themes: first, we discuss the reasoning strategies students use, followed by discussing the findings pertaining to the concepts and features common across reaction types (i.e., the electron-pushing formalism, nucleophiles and electrophiles, acid–base theories, resonance, carbocations, and leaving groups).

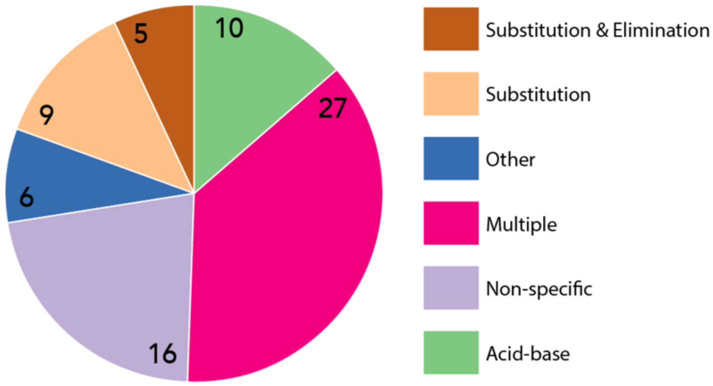


Figure 1.3 Reaction types focused on within the reviewed articles. Reaction types included in the "other" category are alkene addition reactions, oxidation/reduction reactions, acyl transfer reactions, and reactions of aromatic compounds. Articles that included multiple reaction types had tasks that included a range of reactions, with findings that did not focus on students' reasoning with a specific type of reaction.

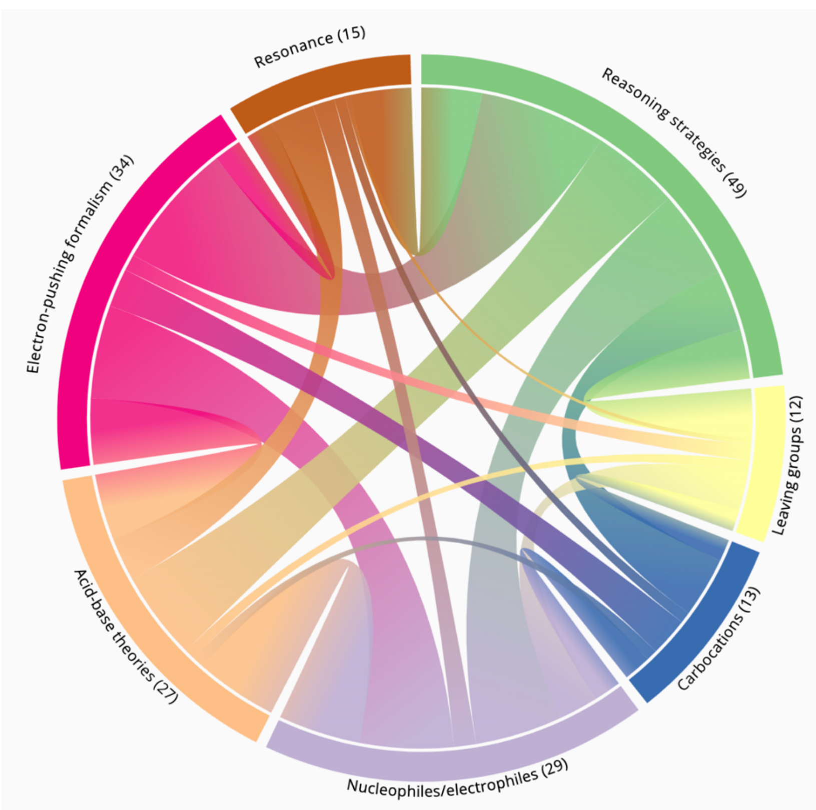


Figure 1.4 Proportion of the reviewed articles with findings pertaining to each theme, with chords indicating the overlap of themes within articles.

1.7 Discussion

The discussion is organized in alignment with the research questions. To address the first research question, we provide an in-depth synthesis of the reasoning strategies reported in the literature for students' descriptions and explanations of reaction mechanisms. This discussion includes an overview of the frameworks researchers use to characterize students' reasoning for organic reaction mechanisms. To address the second research question, we synthesize the findings pertaining to students' reasoning, descriptions, and explanations of concepts and features common to reaction mechanisms (the electron-pushing formalism, nucleophiles and electrophiles, acid–base theories, resonance, carbocations, and leaving groups). Within each section, we aim to describe the findings pertaining to what students can do along with the reported limitations of their understanding. We also include an overview of research related to supporting students' reasoning within each theme. The article concludes with a synthesis of the implications for instruction and future research.

1.7.1 Reasoning strategies for describing and explaining reaction mechanisms

Students use a variety of strategies to describe and explain reaction mechanisms; the most common approaches include teleological reasoning, anthropomorphism, mechanistic reasoning, and causal reasoning. The relationship between these approaches to reasoning and the various frameworks researchers use to characterize reasoning are depicted in Figure 1.5. Students often use teleological reasoning and anthropomorphism when explaining a variety of chemical concepts.[20–25] Teleological reasoning implies that an action occurs to complete a specific goal (e.g., reasoning that a reaction step occurs to form a neutral product).[26,27] Anthropomorphic reasoning gives human-like characteristics to nonhumans (e.g., reasoning that attributes feelings such as wants and needs to atoms or molecules).[26,27] Students from general chemistry through graduate school show a preference for teleological and anthropomorphic explanations over causal explanations, though many students can describe why chemical reactions occur causally.[20,22,25,28–39] Common teleological and anthropomorphic explanations in organic chemistry draw on the octet framework, which students use to suggest that reactions occur in order to complete an octet or because atoms want or need to attain a full octet (i.e., the octet "rule").[24,28,30,40–43] This type of anthropomorphism regarding the "preferences" of atoms has been observed frequently in studies about students' explanations of chemical reactions.[20,22,24,28,35,38,39]

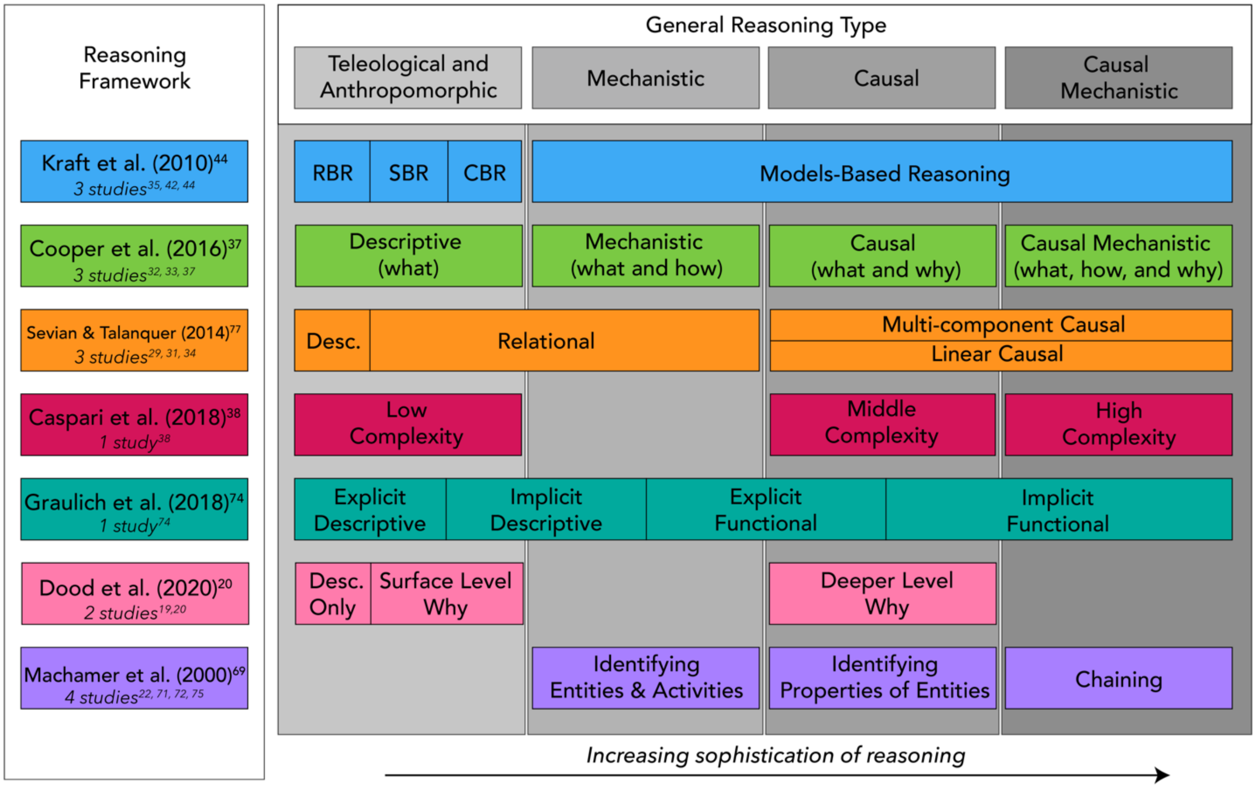


Figure 1.5 Alignment between the general reasoning types identified in the literature and the various frameworks researchers use to characterize students' reasoning. The names for each framework reflect the first instance the framework was presented in the cited literature. Note that many of the frameworks are interrelated and have informed one another, but distinctions are made based on different ways these frameworks have been applied across various studies. Furthermore, the alignment reflected in the figure represents our interpretation of each framework and how they are situated in the literature; we do not intend the alignment of different framework components to be prescriptive, and we recognize additional nuance exists within each framework. The abbreviations within the figure are as follows: for the Kraft et al. (2010)[44] framework, "RBR", "SBR", and "CBR" are "rule-based reasoning", "symbol-based reasoning", and "case-based reasoning", respectively; for the Sevian and Talanquer (2014)[77] and Dood et al. (2020)[20] frameworks, "Desc." is an abbreviation for "Descriptive."

Teleological and anthropomorphic reasoning closely relate to aspects of the reasoning processes framework used to characterize students' reasoning in the organic chemistry education literature (the framework is sometimes referred to as the "modes of reasoning framework", but is referred to as the "reasoning processes framework" to distinguish it from the other modes of reasoning framework).[35,42,44] The framework includes rules-based reasoning (the use of rules, often memorized, to explain phenomena; RBR), case-based reasoning (matching a problem to a similar, memorized case; CBR), and symbol-based reasoning (the manipulation of chemical symbols to reason about phenomena; SBR).[35,44,45] RBR, CBR, and SBR all reflect the surface-level reasoning characteristic of teleological and anthropomorphic reasoning. Of these three, reasoning based on memorized rules is frequently noted in the literature.[23,28–30,34,35,38,40–42,44,46–62] All three reasoning processes contrast with models-based reasoning (MBR), which is the use of a working mental model of a phenomenon that can be applied to dynamic situations (e.g., a mental model for determining the reaction pathway for substitution and elimination reactions).[35,44,63] MBR is similar to models and modeling frameworks,[64–66] which other researchers have used to interpret students' mental models of reaction mechanisms and the related chemical concepts.[30,52,53,55,56,67] Students' more sophisticated mental models are akin to mechanistic and causal reasoning, which are common reasoning strategies described in the literature pertaining to students' explanations of organic chemistry reaction mechanisms.

Researchers in CER define mechanistic reasoning and causal reasoning with slight differences across studies, often depending on the goals of a given analysis; the lack of simple or consensus definitions reflects a survey of faculty which did not reach a consensus definition for mechanistic reasoning.[6] Nevertheless, in a broad sense, mechanistic reasoning encompasses students' descriptions of *how* a reaction occurs, typically at a level lower than the observed phenomena: that is, descriptions of how reactions between molecules proceed through electron movements and changes in bonding.[29,39,68,69] Causal reasoning encompasses students' explanations of *why* a reaction occurs, typically using the chemical or physical properties of the reacting materials to provide explanation that links causes to effects.[32,37,39] Both of these reasoning types typically build upon students' descriptions of *what* occurs in a reaction: describing the starting materials, intermediates, and products. In addition to considering mechanistic reasoning and causal reasoning separately, some researchers define causal mechanistic reasoning as students' combination of both reasoning types to describe *how* and *why* reactions occur.[29,32,33,37]

Several publications address the importance of mechanistic, causal, and causal mechanistic reasoning in the teaching and learning of organic chemistry.[19,20,22,28,29,31–34,37–39,44,70–76] Researchers have used various frameworks to explore these types of reasoning across the articles incorporated in this review; these include the modes of reasoning framework,[29,31,34] frameworks focused on explicit *versus* implicit properties,[19,20,38,59,61,74] and frameworks based in the philosophy of science literature.[22,71,72,75] The modes of reasoning framework, based on the chemical thinking learning progression,[77] characterizes students' reasoning into four modes: descriptive (describing explicit features without further explanation), relational (relating explicit and implicit features as correlations rather than causally), linear causal (relational reasoning that incorporates cause-and-effect), and multicomponent causal (causal reasoning that considers more than one factor).[29,31,34] A key component of this framework is the focus on whether students consider explicit or implicit features. Explicit features are visually present in a representation (e.g., atoms, connectivity, and formal charges), while implicit features must be inferred from the explicit features (e.g., partial charges, electron density, polarizability).

Four other frameworks similarly focus on explicit and implicit features. The first framework, representation mapping, captures the degree to which students base their reasoning on explicit or implicit features (called "abstractness") and whether students partially or strictly match their problem solving to prior problems (called "abstracting").[59,61] The other three frameworks all characterize students' reasoning based on the level of complexity or sophistication when reasoning with explicit and/or implicit features. The first of these characterizes students' mechanistic reasoning when comparing between two similar mechanistic steps, grouping students' reasoning into low, middle, or high complexity of relations.[38] Relations with low complexity describe explicit structural differences as a cause for change; relations with middle complexity describe implicit properties inferred from explicit structural differences as a cause for change; and relations with high complexity describe implicit properties inferred from explicit structural differences as having an electronic effect on change.[38] The second framework, the levels of elaboration framework, focuses on explicit *versus* implicit properties but also considers descriptive *versus* functional elaborations; descriptive elaborations simply state a property while functional elaborations place the property in the context of the reaction.[74] The third framework groups students' reasoning into three levels of sophistication: descriptive only (describing what happens without explaining why), surface level why (describing what happens, using only explicit features or memorized rules to explain why), and deeper why (describing what happens, using implicit features to explain why).[19,20] The modes of reasoning framework and the latter three frameworks focused on explicit *versus* implicit features each characterize students' reasoning on a range of descriptive or surface-level reasoning (akin to teleological and anthropomorphic reasoning) to more complex reasoning (akin to mechanistic and causal reasoning).

The last set of frameworks found in the literature are the philosophy of science frameworks, which are used to more specifically examine aspects of students' mechanistic and causal reasoning.[22,71,72,75] These frameworks are derived from generalized descriptions of mechanisms in the philosophy of science literature.[68,69,78] The frameworks conceptualize mechanisms as composed of *entities* and *activities*, which are defined as the components involved in a mechanism and the actions between the components that produce change, respectively.[69] In the context of organic chemistry reaction mechanisms, entities are electrons, atoms, and molecules while activities are the movement of electrons or molecules that result in the breaking and forming of bonds. Expanded accounts of mechanisms also incorporate the idea of chaining, which is defined as "reasoning about one part of a mechanism on the basis of what is known or conjectured about other parts of the mechanism" (ref 78, p. 362). Chaining can either be forward or backward, where forward chaining is using information about prior mechanistic steps to reason about subsequent steps and backward chaining is reasoning about the prior mechanistic steps using information about the subsequent steps. A framework for discourse analysis developed to identify students' mechanistic reasoning includes (1) describing the target phenomenon, (2) identifying the setup conditions required for the mechanism, (3) identifying the entities, (4) identifying the activities, (5) identifying the properties of entities, (6) identifying the organization of entities, and (7) chaining.[68] By using these frameworks, researchers investigate the complex process of students' reasoning for tasks eliciting written and verbal explanations of reaction mechanisms.[22,71,72,75]

Both mechanistic and causal reasoning have been described as more "scientific" than other types of reasoning due to their relation to the process of scientific inquiry.[68] Indeed, "scientific explanation" is named as one of eight scientific practices in the Framework for K-12 Science Education,[79] which specifically mentions that scientific explanations "describe the mechanisms that support cause and effect inferences about them" (ref 79, p. 67). Studies indicate that students can engage in causal reasoning and use mechanistic reasoning to describe reactions, though students need further support to engage in multivariate reasoning.[22,25,28,29,31–39,44,49,51,58,70–72,75,76] Notably, a variety of studies indicate that students' reasoning with mechanisms is influenced by the task and students' familiarity with the problem.[23,30–32,36–38,42,44,73,80,81] These findings indicate that instructors should carefully design problems that elicit reasoning aligned with the learning objectives for the course. Supporting students' reasoning is important because their performance on organic chemistry tasks can correlate with the use of causal and mechanistic reasoning; specifically promoting causal and mechanistic reasoning when using the EPF to solve unfamiliar problems is a beneficial practice.[21,31–33,36,37,44,73,80] Redesigned curricula[21,32,33,36,37,80,82–84] and carefully designed instructional materials, such as constructed response items,[17,20,29,32,33,37] case comparisons,[29,38,51,70,74] animations or videos,[30,42,76] and writing-to-learn assignments,[72] all show promise for promoting students' mechanistic and causal reasoning.

1.7.2 Concepts and features common across reaction types

The reviewed articles focus on how students describe and explain a variety of reaction mechanisms in organic chemistry. From these studies, there are various findings about students' reasoning with common concepts and features across reaction mechanisms. In general, these studies focus on undergraduate students; unless otherwise noted, the presented findings pertain to this population. The following sections synthesize the findings pertaining to the most prevalent concepts or features identified in the review:

1. The electron-pushing formalism
2. Nucleophiles and electrophiles
3. Acid–base theories
4. Resonance
5. Carbocations
6. Leaving groups

1.7.3 The electron-pushing formalism

The EPF is central to describing and explaining mechanisms in the practice of organic chemistry, and many studies have focused on students' understanding of the EPF and their descriptions of electron movement. The ability to explain mechanisms is fundamental to organic chemistry,[6,85] and studies indicate that the ability to propose mechanisms and compare reactivity is predictive of students' success both in a nonmajors course and when solving unfamiliar mechanism problems.[73,80,86] However, few students consider patterns of electron movement when sorting reactions, though professors and doctoral students primarily organize reactions by mechanism type.[62,87–89] Additionally, both undergraduate and graduate students can correctly predict the product without using the EPF or can reproduce the EPF for specific reactions from memory, often without being able to explain the underlying chemical concepts.[46,49,50,59,60,73,80,90,91] Furthermore, when students are unable to recall a mechanism from memory or face an unfamiliar mechanism, they struggle to predict the products using conceptual understanding.[59,73,90,92]

Further evidence suggests that many students at both the undergraduate and graduate levels may struggle to make the connection between the notation of the EPF and the electron movement it represents.[30,32,36,46–48,52,53,60,67,90,91] In one study, graduate students focused primarily on structural changes when describing reaction mechanisms, viewing functional groups as book-keeping devices and the EPF arrows themselves, rather than the electrons they represent, as agents for change.[67] Similar studies have noted the apparent "meaninglessness" of the EPF for students (including graduate students), where limited understanding of fundamental concepts might prevent students from recognizing the usefulness of the EPF for predicting and explaining reaction outcomes.[47,48,90] When using the EPF to solve problems, students often take a teleological or product-oriented approach, such as depicting mechanisms to access infeasible charged species to react in further steps.[23,35,47,58,90,93] Together, the findings indicate that students may not be thinking about reactions mechanistically or understanding reactions conceptually.

In contrast to the findings indicating that students do not associate the EPF with physical meaning, other findings indicate that students are able to describe and explain electron movement and depict the flow of electrons from electron-rich to electron-poor atoms.[21,23,30,32,33,35–37,42,59,72,80,91,93,94] Students often exhibit understanding of electron movement when providing verbal and written explanations situated within peer interactions.[35,46,72,94] Other studies suggest the value of supporting students' mechanistic reasoning through learning modules focused on the EPF, teaching with animated models of reaction mechanisms, or encouraging students to work through mechanisms with a mobile device application that offers guidance and immediate feedback.[42,52,53,95] Furthermore, a number of studies focus specifically on how students in revised organic chemistry curricula,[82,84] or who participated in a revised general chemistry curriculum,[83] engage with the EPF and use mechanistic reasoning, both when prompted to draw mechanisms with the EPF and when not explicitly told to do so.[21,30,32,34,36,37,42,80,89] The transformed curricula in these studies place more focus on mechanistic reasoning and using the EPF.[82,84] These findings suggest the EPF is useful for guiding students to reason about reaction mechanisms at the electronic level when presented with additional instructional emphasis.

1.7.4 Nucleophiles and electrophiles

A number of studies have probed how students define and understand nucleophiles and electrophiles. Students often correctly identify, define, and provide examples of nucleophiles and electrophiles,[93,96,97] a skill that predicted students' overall success in a study focused on a nonmajors organic chemistry course.[86] However, while students correctly use the terms, both undergraduate and graduate students may struggle to describe nucleophiles and electrophiles beyond explicit structural features or to use them operationally.[46,49,51,67,74,93,97,98] For example, many students use formal charges (an explicit feature) as a primary reason for identifying nucleophiles and electrophiles, rarely mentioning implicit features such as polarizability, partial charges, inductive effects, and resonance.[21,51,72,93,98] Similarly, when proposing mechanisms, some students use teleological reasoning to form a positive charge in order for a nucleophilic addition to occur.[21,22] Other students feel they need to know the reaction mechanism to determine the nucleophile and electrophile.[93,98] Some studies have indicated that students especially face challenges with understanding electrophiles.[50,93,97] Challenges with electrophiles may be due to instructional emphasis on nucleophiles, which leaves electrophiles as "simply the other starting material in the reaction" (ref 93, p. 800). This is reflected in how students often describe reactions with the nucleophile doing the action, which places emphasis on the nucleophile.[72,94]

Altogether, the findings indicate a surface-level understanding of the relationship between structure and function and the possibility of students identifying nucleophiles and electrophiles based on memorized, explicit features rather than implicit properties. This is supported by findings that indicate challenges undergraduate and graduate students face with connecting nucleophiles and electrophiles to concepts such as Lewis acid–base theory.[23,25,47,49,51,67,72,90,93,96] Additionally, findings from one study suggested students' use of the terms "nucleophile" and "electrophile" was not related to causal mechanistic reasoning,[32] though students in another study who elaborated on nucleophilicity performed better when comparing reactivity between molecules.[74] While students often have knowledge of the concepts and definitions, they still face challenges with integrating that knowledge with other concepts or into their reasoning about reaction mechanisms.

It is necessary for instructors and researchers to support students' understanding of nucleophiles and electrophiles, as practicing chemists consider nucleophilicity and electrophilicity seamlessly when considering reaction mechanisms while students in organic chemistry laboratory courses often do not.[99] To address this, studies on instructional strategies demonstrate approaches for supporting students' understanding of nucleophilicity and electrophilicity and using these concepts in their reasoning. These include a POGIL activity,[100] case comparison problems,[51] and adaptive interventions based on automated text analysis of students' written explanations of mechanisms.[16,19,20] Furthermore, studies taking place within a revised, mechanisms-first curriculum[84] have shown that students consider implicit properties including partial charges, electronegativity, and reactivity when reasoning about nucleophiles and electrophiles (though students rarely consider electron density or orbitals).[36,87–89] These studies indicate the value of research-informed instructional strategies to support students' functional understanding of nucleophiles and electrophiles.

1.7.5 Acid–base theories

Students' understanding of acid–base concepts is relevant to their understanding of reaction mechanisms more broadly, as many reaction mechanisms contain at least one acid–base step.[6,85,101] Studies indicate that some undergraduate and graduate students focus on the surface features of molecules (rather than interpreting the implicit properties or functions) and may not recognize species as acids or bases.[47,48,53,55–57,67,102] However, the ability to recognize acids and bases is important for students' success in organic chemistry.[86] Students hold multiple models of acids and bases which align with the Arrhenius, Brønsted–Lowry, and Lewis theories, though they may conflate definitions across theories, have difficulties categorizing acids and bases as strong or weak, or face challenges with articulating their reasoning for acid–base tasks depending on the theory being used.[31,55–57,67,97,102] However, students are able to use causal reasoning to connect concepts such as pK_a values, conjugate acid strength, base strength, and the direction of acid–base equilibria.[31,53]

The theories students use in their reasoning is related to their understanding of mechanisms. Specifically, Lewis acid–base theory aligns with mechanistic reasoning by focusing on the underlying movement of electrons, as noted in a survey of organic chemistry faculty.[6] Studies indicate that when students use Lewis acid–base theory in explanations of acid–base reactions, their reasoning is more sophisticated, associated with better performance on assessment items, and associated with correct use of the electron-pushing formalism.[17,33,37] However, some studies found that students were more likely to consider Brønsted–Lowry theory relative to Lewis theory, even when the Brønsted–Lowry definitions of acids and bases were less applicable to the task.[53,72,96,97,103] Acid–base theories also closely relate to other concepts necessary for mechanistic and causal reasoning, which poses additional challenges for students. For instance, conflating acidity and basicity with the concepts of electrophilicity and nucleophilicity is a well-documented challenge for students, including graduate students.[23,25,47,49,51,67,72,90,93,96] Other studies indicate that students have challenges with making appropriate connections between acid–base concepts and nucleophilicity and electrophilicity.[72,96,98]

To support students' understanding of acid–base theories in the organic chemistry context, researchers have called for scaffolding students' reasoning with acid–base theories to support consideration of Lewis theory and the associated implicit properties such as polarity and partial charges.[37,53] Studies demonstrate that a revised general chemistry curriculum which emphasizes models-based reasoning[83] may support students' causal and mechanistic reasoning with Lewis acid–base theory in organic chemistry.[33,37] Other research to support students' reasoning with acid–base theories includes writing-to-learn assignments[103] and adaptive tutorials that use machine learning models to identify students' use of Lewis acid–base theory in order to promote understanding of the Lewis acid–base model for proton-transfer reactions or reaction steps.[16–19]

1.7.6 Resonance

Students' understanding of resonance in the context of organic chemistry reaction mechanisms has been the focus of a number of studies. Both undergraduate and graduate students often consider resonance in their reasoning, occasionally exhibiting over-reliance on the concept.[47,52,53,58,97] When reasoning about resonance, students exhibit varied interpretations of how it influences reactivity and often apply the concept to inappropriate species when determining mechanistic steps.[51–54] For example, students often know to consider resonance when determining relative acid strengths, but some students may not understand the relationship conceptually or focus on the capabilities for resonance in the acid itself rather than the conjugate base.[53–57,102] Additionally, students occasionally draw resonance contributors as intermediates when drawing mechanisms, representing resonance structures as discrete species formed during the reaction.[21]

Resonance is challenging for students to understand because of the limitations of two-dimensional structural representations and how they relate to the meaning behind the concept. One common misunderstanding is that resonance structures represent multiple, distinct forms that alternate in an equilibrium process;[41,52,53,104] some researchers suggest referring to resonance as "delocalization" to address this misunderstanding.[105] Students also exhibit difficulties describing the resonance hybrid, commonly conflating major contributors with the resonance hybrid.[41] Students' challenges with the representational meaning of resonance relates to the finding that students' abilities for drawing resonance structures may not be associated with their conceptual understanding.[41,106] Similarly, students' understanding of resonance tends to focus more on drawing resonance structures through the movement of electrons and bonds rather than the conceptual meaning.[41,54] This is reflected by another finding that, when solving problems, students consider resonance when cued by the presence of nonbonding electron pairs but may not consider resonance when nonbonding electron pairs are not explicitly shown.[51–53]

To support students' conceptual understanding of resonance, researchers suggest instruction that uses analogies and focuses on the limitations of chemical representations, rather than focusing on the rules and guidelines for drawing resonance structures.[41,54] One study describes a writing-to-learn assignment that affords students the opportunity to explore the conceptual aspects of resonance.[54] Other studies report a mobile device application that provided hints which supported students in appropriately considering the influence of resonance on reactivity, suggesting the value of providing students additional guidance for how to apply resonance when solving mechanism problems.[52,53] Additionally, researchers have described ten essential learning outcomes for resonance that can guide instruction and assessment on the concept.[105]

1.7.7 Carbocations

Various studies examine students' understanding of carbocations, which are important for many reactions taught in introductory organic chemistry (e.g., S_N1, E1, and addition reactions). Although carbocation intermediates are common, studies indicate that students often do not recognize or depict carbocation intermediates when appropriate.[49,91,92] Furthermore, when describing the formation of carbocation intermediates, students often use teleological reasoning (e.g., reasoning that a leaving group departs so a stable tertiary carbocation can be formed).[20] Students commonly invoke the trend that more substituted carbocations are more stable as the sole reason for determining carbocation stability, often without providing chemical reasoning (such as hyperconjugative effects); students occasionally misremember the trend and incorrectly determine carbocation stability.[20,21,29,38,50,52,81] However, students can exhibit deeper understanding of influences on carbocation stability, such as hyperconjugation or induction.[35,74]

Researchers have described various approaches to support students' chemical reasoning about carbocation stability beyond the number-of-substituents argument; these approaches have elicited students' reasoning with properties including the size of substituents, resonance, partial charges, and inductive effects.[19,20,38,52,70] Problems to support students' reasoning about carbocation stability include case comparisons, which elicit students' reasoning about implicit properties, especially when paired with an instructional scaffold.[38,70] Another study reported a mobile device application in which students worked through an addition mechanism that required them to consider resonance stabilization rather than substitution to correctly determine carbocation stability; the application guided students to consider resonance and the alternative mechanistic pathway if their initial response formed the minor product.[52] Other research to support students' reasoning about carbocations includes an adaptive tutorial which uses lexical analysis to provide targeted feedback to help students develop reasoning about electron density when deciding carbocation stability for S_N1 reactions.[19,20]

1.7.8 Leaving groups

A variety of studies address students' understanding of leaving groups in organic chemistry. Students often use leaving groups to categorize reactions and can correctly identify leaving groups when provided with mechanistic steps;[40,87,107] however, students can struggle to identify "hidden" leaving groups in other reaction types (e.g., acetal reactions).[89] Many students refer to leaving groups as "good" without providing explanation, suggesting that students might be using surface features to memorize leaving group ability rather than using chemical reasoning.[20,40,87] Similarly, students' reasoning about leaving group departure steps is often teleological rather than causal; for example, students often focus on creating good leaving groups rather than reasoning based on chemical properties.[22,23,61] However, some students invoke the octet rule, charge to size ratio, and electronegativity to reason about leaving group ability and leaving group departure steps.[38,40,70]

Studies have explored how the nature of mechanism tasks influences students' reasoning about leaving groups. For instance, students' reasoning when proposing mechanisms with leaving group departure steps may be more teleological, rather than based on chemical properties, when the product of a reaction is shown.[23] Additionally, problems that allow students to rely on surface features rather than the implicit properties of leaving groups can mask student difficulties with implicit features, as it allows students to correctly answer questions without using deeper reasoning.[74] Studies indicate that contrasting cases encourage students to consider implicit properties when reasoning about leaving group departure steps;[38,70] when paired with an instructional scaffold, contrasting cases encourage students to consider multiple implicit influences.[70] Other researchers have used lexical analysis of students' responses to constructed response items to develop adaptive tutorials that promote students' consideration of leaving group ability based on chemical properties.[19,20]

1.8 Conclusions and implications

The reviewed literature provides thorough detail of students' various approaches to reasoning about reaction mechanisms in organic chemistry (e.g., teleological *versus* causal reasoning) and how students understand and reason with concepts and features common across reaction types (i.e., the EPF, nucleophiles and electrophiles, acid–base theories, resonance, carbocations, and leaving groups). The central idea of the reviewed literature is that students can explain mechanisms using a range of reasoning approaches, and that the range in students' reasoning is reflected in their demonstrated understanding of fundamental organic chemistry concepts. Students must integrate their understanding of fundamental concepts into their reasoning with mechanisms, which is central to learning organic chemistry. From the reviewed studies, there are several implications for both instruction and future research.

1.8.1 Implications for instruction

The synthesized findings in each section of this review provide valuable details about the range of students' understanding, from common misunderstandings of specific topics to students' abilities to demonstrate sophisticated reasoning. Instructors can use this information to better perceive the range of students' understanding that may be present in their classrooms. The primary implications for instruction are (1) that supporting students' reasoning requires providing students with the opportunity to practice explaining why reactions happen, and (2) to construct assessments which emphasize mechanistic and causal reasoning. To do this, instructors can explicitly ask students to provide their reasoning for mechanisms alongside mechanistic arrows to help students develop causal mechanistic reasoning skills.[108] In order to support students' success with these tasks, instruction should emphasize why electrons move the way they do, including a focus on building connections to fundamental concepts for explaining why chemical phenomena occur.

The call to provide students with more opportunities to develop explanations for why reactions occur is accompanied by the caveat that such opportunities must be carefully designed. A variety of studies indicate that the nature of a task can influence students' reasoning; for example, the amount of information given to students, such as whether or not the product of a reaction is shown, can influence whether students reason teleologically or causally.[23,31,32,37,81] Smaller task features, such as minor changes in wording or whether lone electron pairs are shown on Lewis structures, can also influence students' reasoning. Instructors must carefully design class activities or problem sets to help students think critically about the features present in reactions and which features are relevant for given situations.

The large body of research demonstrates students' challenges with mechanisms across different contexts and experience levels; this provides confirmation that students would benefit from instruction that supports causal and mechanistic reasoning. The reviewed articles specify a variety of instructional strategies to support this learning outcome. Curricular reform is the primary evidence-based strategy for generally supporting students' reasoning.[21,30,32,33,36,37,42,80,82–84,87–89] Research also demonstrates the value of other instructional strategies which may be more feasible to implement; these include the following:

- Teaching with animated models or videos of reactions.[30,42,76]
- Learning modules and adaptive interventions.[16,19,95]
- Constructed response items requiring students to articulate their reasoning.[17,20,29,32,33,37,108]
- Case comparison problems with instructional scaffolding to support students' consideration of implicit properties.[29,38,51,70,74,109,110]
- Tasks involving peer interactions, including writing-to-learn assignments with peer review.[35,46,54,72,94,100,103,111]
- A mobile device application for solving mechanism problems that provides immediate feedback.[52,53]

These approaches can promote students' engagement with understanding why reactions occur, rather than allowing students to rely on rote memorization or surface-level reasoning.

1.8.2 Implications for research

This review provides an overview of the large body of research exploring students' reasoning with mechanisms, providing insight into the avenues for future research. First, further research is needed to more deeply understand the teaching and learning of the concepts and features necessary for reasoning with mechanisms. Existing studies identify the learning outcomes for acid–base chemistry and resonance in the context of introductory organic chemistry;[101,105] an avenue for future research includes similar studies focused on the learning outcomes for other topics such as the electron-pushing formalism and nucleophiles and electrophiles. Additionally, much of the existing literature focuses on students' reasoning with acid–base reactions and substitution reactions; further research focusing on other common reaction types (such as alkene addition reactions, reactions of aromatic compounds, reactions of carbonyl compounds, and radical reactions) would provide useful insight into how students reason with mechanisms. Existing studies also largely focus on undergraduate students in introductory organic chemistry courses, where more research on other populations (such as advanced undergraduates, graduate students, faculty, and practicing chemists) would provide deeper insight into the range of reasoning approaches and how reasoning develops with experience. Another necessary area of research is to explore how instructors learn to teach mechanisms in organic chemistry, which can inform graduate education and professional development for instructors.

An important consideration for further research is recognizing that much of the existing literature in the area is framed through a deficit approach, focusing on what students cannot do as opposed to what they can do. While some of the reported findings indicate students' successes, a general direction for future research is to explore what students can do and what instructional strategies can better support student learning. Specifically, there is a need for further research on the instructional strategies already reported in the literature (summarized in the Implications for Instruction section); future studies should more deeply investigate how and why these strategies work, how they can be adjusted for different contexts, whether they support some students more than others, and whether the learning outcomes supported by these strategies are lasting. These studies could leverage quantitative techniques and longitudinal, cross-sectional, or experimental methodologies, in addition to the qualitative methodologies that are prevalent in the existing literature. There is also value for future research on novel instructional strategies that promote students' reasoning with mechanisms through tasks reflective of the work of practicing organic chemists.[108,109,112] These strategies could be derived from additional future research exploring faculty and practicing chemists' reasoning with mechanisms. Such research to support instruction should build on the synthesized body of literature that highlights the range of students' understanding and reasoning with reaction mechanisms.

1.9 Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would additionally like to acknowledge Jeffrey Raker, Ginger Shultz, and Solaire Finkenstaedt-Quinn for their support and discussions regarding this manuscript.

1.10 References

(1) Goodwin, W. Mechanisms and Chemical Reaction. In *Philosophy of Chemistry*; Hendry, R., Needham, P., Woody, A., Eds.; Handbook of the Philosophy of Science; North Holland, 2012; pp 309–327.

(2) Goodwin, W. Explanation in Organic Chemistry. *Annals of the New York Academy of Sciences* **2003**, *988* (1), 141–153. https://doi.org/10.1111/j.1749-6632.2003.tb06093.x.

(3) Morrison, R. T.; Boyd, R. N. *Organic Chemistry*; Allyn & Bacon: Boston, MA, 1959.

(4) Grossman, R. B. *The Art of Writing Reasonable Organic Reaction Mechanisms*, 2nd ed.; Springer Verlag: New York, 2003.

(5) Wheeler, D. M. S.; Wheeler, M. M. Trends in the Teaching of Organic Chemistry: A Survey of Some Textbooks. *J. Chem. Educ.* **1982**, *59* (10), 863. https://doi.org/10.1021/ed059p863.

(6) Bhattacharyya, G. From Source to Sink: Mechanistic Reasoning Using the Electron-Pushing Formalism. *Journal of Chemical Education* **2013**, *90*, 1282–1289. https://doi.org/dx.doi.org/10.1021/ed300765k.

(7) Johnstone, A. H. Why Is Science Difficult to Learn? Things Are Seldom What They Seem. *J Comp Assist Learn* **1991**, *7* (2), 75–83. https://doi.org/10.1111/j.1365-2729.1991.tb00230.x.

(8) Graulich, N. The Tip of the Iceberg in Organic Chemistry Classes: How Do Students Deal with the Invisible? *Chem. Educ. Res. Pract.* **2015**, *16* (1), 9–21. https://doi.org/10.1039/C4RP00165F.

(9) Colquhoun, H. L.; Levac, D.; O'Brien, K. K.; Straus, S.; Tricco, A. C.; Perrier, L.; Kastner, M.; Moher, D. Scoping Reviews: Time for Clarity in Definition, Methods, and Reporting.

Journal of Clinical Epidemiology **2014**, *67* (12), 1291–1294. https://doi.org/10.1016/j.jclinepi.2014.03.013.

(10) Levac, D.; Colquhoun, H.; O'Brien, K. K. Scoping Studies: Advancing the Methodology. *Implementation Sci* **2010**, *5* (1), 69. https://doi.org/10.1186/1748-5908-5-69.

(11) Pham, M. T.; Rajić, A.; Greig, J. D.; Sargeant, J. M.; Papadopoulos, A.; McEwen, S. A. A Scoping Review of Scoping Reviews: Advancing the Approach and Enhancing the Consistency. *Res. Syn. Meth.* **2014**, *5* (4), 371–385. https://doi.org/10.1002/jrsm.1123.

(12) Munn, Z.; Peters, M. D. J.; Stern, C.; Tufanaru, C.; McArthur, A.; Aromataris, E. Systematic Review or Scoping Review? Guidance for Authors When Choosing between a Systematic or Scoping Review Approach. *BMC Med Res Methodol* **2018**, *18* (1), 143. https://doi.org/10.1186/s12874-018-0611-x.

(13) Lincoln, Y. S.; Guba, E. G. *Naturalistic Inquiry*; Sage Publications: Beverly Hills, California, 1985.

(14) Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* **2006**, *3* (2), 77–101. https://doi.org/10.1191/1478088706qp063oa.

(15) Watts, F. M.; Finkenstaedt-Quinn, S. A. The Current State of Methods for Establishing Reliability in Qualitative Chemistry Education Research Articles. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 565–578. https://doi.org/10.1039/D1RP00007A.

(16) Dood, A. J.; Fields, K. B.; Cruz-Ramírez de Arellano, D.; Raker, J. R. Development and Evaluation of a Lewis Acid–Base Tutorial for Use in Postsecondary Organic Chemistry Courses. *Can. J. Chem.* **2019**, *97* (10), 711–721. https://doi.org/10.1139/cjc-2018-0479.

(17) Dood, A. J.; Fields, K. B.; Raker, J. R. Using Lexical Analysis To Predict Lewis Acid–Base Model Use in Responses to an Acid–Base Proton-Transfer Reaction. *J. Chem. Educ.* **2018**, *95* (8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177.

(18) Yik, B. J.; Dood, A. J.; Cruz-Ramírez de Arellano, D.; Fields, K. B.; Raker, J. R. Development of a Machine Learning-Based Tool to Evaluate Correct Lewis Acid–Base Model Use in Written Responses to Open-Ended Formative Assessment Items. *Chem. Educ. Res. Pract.* **2021**, *22* (4), 866–885. https://doi.org/10.1039/D1RP00111F.

(19) Dood, A. J.; Dood, J. C.; Cruz-Ramírez de Arellano, D.; Fields, K. B.; Raker, J. R. Using the Research Literature to Develop an Adaptive Intervention to Improve Student Explanations

of an S_N1 Reaction Mechanism. *J. Chem. Educ.* **2020**, *97* (10), 3551–3562. https://doi.org/10.1021/acs.jchemed.0c00569.

(20) Dood, A. J.; Dood, J. C.; Cruz-Ramírez de Arellano, D.; Fields, K. B.; Raker, J. R. Analyzing Explanations of Substitution Reactions Using Lexical Analysis and Logistic Regression Techniques. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 267–286. https://doi.org/10.1039/C9RP00148D.

(21) Galloway, K. R.; Stoyanovich, C.; Flynn, A. B. Students' Interpretations of Mechanistic Language in Organic Chemistry before Learning Reactions. *Chem. Educ. Res. Pract.* **2017**, *18* (2), 353–374. https://doi.org/10.1039/C6RP00231E.

(22) Caspari, I.; Weinrich, M. L.; Sevian, H.; Graulich, N. This Mechanistic Step Is "*Productive*": Organic Chemistry Students' Backward-Oriented Reasoning. *Chem. Educ. Res. Pract.* **2018**, *19* (1), 42–59. https://doi.org/10.1039/C7RP00124J.

(23) DeCocq, V.; Bhattacharyya, G. TMI (Too Much Information)! Effects of given Information on Organic Chemistry Students' Approaches to Solving Mechanism Tasks. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 213–228. https://doi.org/10.1039/C8RP00214B.

(24) Nicoll, G. A Report of Undergraduates' Bonding Misconceptions. *International Journal of Science Education* **2001**, *23* (7), 707–730. https://doi.org/10.1080/09500690010025012.

(25) Bhattacharyya, G. Trials and Tribulations: Student Approaches and Difficulties with Proposing Mechanisms Using the Electron-Pushing Formalism. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 594–609. https://doi.org/10.1039/C3RP00127J.

(26) Talanquer, V. Explanations and Teleology in Chemistry Education. *International Journal of Science Education* **2007**, *29* (7), 853–870. https://doi.org/10.1080/09500690601087632.

(27) Talanquer, V. When Atoms Want. *J. Chem. Educ.* **2013**, *90* (11), 1419–1424. https://doi.org/10.1021/ed400311x.

(28) Weinrich, M. L.; 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.

(29) Bode, 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. *Journal of Chemical Education* **2019**, *96*, 1068–1082. https://doi.org/10.1021/acs.jchemed.8b00719.

(30) Bongers, A.; Northoff, G.; Flynn, A. B. Working with Mental Models to Learn and Visualize a New Reaction Mechanism. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 554–569. https://doi.org/10.1039/C9RP00060G.

(31) Deng, J. M.; Flynn, A. B. Reasoning, Granularity, and Comparisons in Students' Arguments on Two Organic Chemistry Items. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 749–771. https://doi.org/10.1039/D0RP00320D.

(32) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *Journal of Chemical Education* **2020**, *9*, 313–327.

(33) Crandell, O. M.; Kouyoumdjian, H.; Underwood, S. M.; Cooper, M. M. Reasoning about Reactions in Organic Chemistry: Starting It in General Chemistry. *Journal of Chemical Education* **2019**, *96* (2), 213–226. https://doi.org/10.1021/acs.jchemed.8b00784.

(34) Weinrich, M. L.; Talanquer, V. Mapping Students' Modes of Reasoning When Thinking about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 394–406. https://doi.org/10.1039/C5RP00208G.

(35) Christian, K.; Talanquer, V. Modes of Reasoning in Self-Initiated Study Groups in Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 286–295. https://doi.org/10.1039/C2RP20010D.

(36) Webber, D. M.; Flynn, A. B. How Are Students Solving Familiar and Unfamiliar Organic Chemistry Mechanism Questions in a New Curriculum? *J. Chem. Educ.* **2018**, *95* (9), 1451–1467. https://doi.org/10.1021/acs.jchemed.8b00158.

(37) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid−Base Reactions. *Journal of Chemical Education* **2016**, *93*, 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(38) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/C8RP00131F.

(39) Yan, F.; Talanquer, V. Students' Ideas about How and Why Chemical Reactions Happen: Mapping the Conceptual Landscape. *International Journal of Science Education* **2015**, *37* (18), 3066–3092.

(40) Popova, M.; Bretz, S. L. Organic Chemistry Students' Understandings of What Makes a Good Leaving Group. *Journal of Chemical Education* **2018**, *95*, 1094–1101. https://doi.org/10.1021/acs.jchemed.8b00198.

(41) 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.

(42) Bongers, A.; Beauvoir, B.; Streja, N.; Northoff, G.; Flynn, A. B. Building Mental Models of a Reaction Mechanism: The Influence of Static and Animated Representations, Prior Knowledge, and Spatial Ability. *Chem. Educ. Res. Pract.* **2020**, *21* (2), 496–512. https://doi.org/10.1039/C9RP00198K.

(43) Taber, K. S. A Common Core to Chemical Conceptions: Learners' Conceptions of Chemical Stability, Change and Bonding. In *Concepts of Matter in Science Education*; Tsaparlis, G., Sevian, H., Eds.; Innovations in Science Education and Technology; Springer Netherlands: Dordrecht, 2013; Vol. 19, pp 391–418. https://doi.org/10.1007/978-94-007-5914-5_19.

(44) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable Reasoning: Multi-Variate Problem-Solving in Organic Chemistry. *Chemistry Education Research and Practice* **2010**, *11*, 281–292. https://doi.org/10.1039/C0RP90003F.

(45) Kolodner, J. L. An Introduction to Case-Based Reasoning. *Artificial Intelligence Review* **1992**, *6*, 3–34. https://doi.org/10.1007/BF00155578.

(46) Wilson, S. B.; Varma-Nelson, P. Characterization of First-Semester Organic Chemistry Peer-Led Team Learning and Cyber Peer-Led Team Learning Students' Use and Explanation of Electron-Pushing Formalism. *J. Chem. Educ.* **2019**, *96* (1), 25–34. https://doi.org/10.1021/acs.jchemed.8b00387.

(47) Ferguson, R.; Bodner, G. M. Making Sense of the Arrow-Pushing Formalism among Chemistry Majors Enrolled in Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 102–113. https://doi.org/10.1039/B806225K.

(48) Anderson, T. L.; Bodner, G. M. What Can We Do about 'Parker'? A Case Study of a Good Student Who Didn't 'Get' Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–101. https://doi.org/10.1039/B806223B.

(49) Cruz-Ramírez de Arellano, D.; Towns, M. H. Students' Understanding of Alkyl Halide Reactions in Undergraduate Organic Chemistry. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 501–515. https://doi.org/10.1039/C3RP00089C.

(50) Sendur, G. An Examination of Pre-Service Chemistry Teachers' Meaningful Understanding and Learning Difficulties about Aromatic Compounds Using a Systemic Assessment Questions Diagram. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 113–140. https://doi.org/10.1039/C9RP00080A.

(51) Watts, F. M.; Zaimi, I.; Kranz, D.; Graulich, N.; Shultz, G. V. Investigating Students' Reasoning over Time for Case Comparisons of Acyl Transfer Reaction Mechanisms. *Chem. Educ. Res. Pract.* **2021**, *22* (2), 364–381. https://doi.org/10.1039/D0RP00298D.

(52) Finkenstaedt-Quinn, S. A.; Watts, F. M.; Petterson, M. N.; Archer, S. R.; Snyder-White, E. P.; Shultz, G. V. Exploring Student Thinking about Addition Reactions. *J. Chem. Educ.* **2020**, *97* (7), 1852–1862. https://doi.org/10.1021/acs.jchemed.0c00141.

(53) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.; Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions *via* App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (3), 878–892. https://doi.org/10.1039/C9RP00260J.

(54) Brandfonbrener, P. B.; Watts, F. M.; Shultz, G. V. Organic Chemistry Students' Written Descriptions and Explanations of Resonance and Its Influence on Reactivity. *J. Chem. Educ.* **2021**, *98* (11), 3431–3441. https://doi.org/10.1021/acs.jchemed.1c00660.

(55) McClary, L.; Talanquer, V. College Chemistry Students' Mental Models of Acids and Acid Strength. *J. Res. Sci. Teach.* **2011**, *48* (4), 396–413. https://doi.org/10.1002/tea.20407.

(56) Bhattacharyya, G. Practitioner Development in Organic Chemistry: How Graduate Students Conceptualize Organic Acids. *Chem. Educ. Res. Pract.* **2006**, *7* (4), 240–247. https://doi.org/10.1039/B5RP90024G.

(57) McClary, L.; Talanquer, V. Heuristic Reasoning in Chemistry: Making Decisions about Acid Strength. *International Journal of Science Education* **2011**, *33* (10), 1433–1454. https://doi.org/10.1080/09500693.2010.528463.

(58) Zotos, E. K.; Tyo, J. J.; Shultz, G. V. University Instructors' Knowledge for Teaching Organic Chemistry Mechanisms. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 715–732. https://doi.org/10.1039/D0RP00300J.

(59) 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.

(60) Graulich, N. Intuitive Judgments Govern Students' Answering Patterns in Multiple-Choice Exercises in Organic Chemistry. *J. Chem. Educ.* **2015**, *92* (2), 205–211. https://doi.org/10.1021/ed500641n.

(61) Sevian, H.; Bernholt, S.; Szteinberg, G. A.; Auguste, S.; Pérez, L. C. Use of Representation Mapping to Capture Abstraction in Problem Solving in Different Courses in Chemistry. *Chem. Educ. Res. Pract.* **2015**, *16* (3), 429–446. https://doi.org/10.1039/C5RP00030K.

(62) Graulich, N.; Bhattacharyya, G. Investigating Students' Similarity Judgments in Organic Chemistry. *Chem. Educ. Res. Pract.* **2017**, *18* (4), 774–784. https://doi.org/10.1039/C7RP00055C.

(63) Nersessian, N. J. Model-Based Reasoning in Conceptual Change. In *Model-Based Reasoning in Scientific Discovery*; Magnani, L., Nersessian, N. J., Thagard, P., Eds.; Springer US: Boston, MA, 1999; pp 5–22. https://doi.org/10.1007/978-1-4615-4813-3_1.

(64) Briggs, M. Models and Modeling. In *Theoretical Frameworks for Research In Chemistry and Science Education*; Bodner, G. M., Orgill, M., Eds.; Pearson Prentice Hall: Upper Saddle River, NJ, 2007; pp 72–85.

(65) Briggs, M.; Bodner, G. A Model of Molecular Visualization. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer Netherlands: Dordrecht, 2005; pp 61–72. https://doi.org/10.1007/1-4020-3613-2_5.

(66) Greca, I. M.; Moreira, M. A. Mental Models, Conceptual Models, and Modelling. *International Journal of Science Education* **2000**, *22* (1), 1–11. https://doi.org/10.1080/095006900289976.

(67) 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.

(68) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/DOI 10.1002/sce.20264.

(69) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25.

(70) Caspari, I.; Graulich, N. Scaffolding the Structure of Organic Chemistry Students' Multivariate Comparative Mechanistic Reasoning. *International Journal of Physics and Chemistry Education* **2019**, *11* (2), 31–43. https://doi.org/10.12973/ijpce/211359.

(71) Keiner, L.; Graulich, N. Transitions between Representational Levels: Characterization of Organic Chemistry Students' Mechanistic Features When Reasoning about Laboratory Work-up Procedures. *Chemistry Education Research and Practice* **2020**, *21*, 469–482. https://doi.org/10.1039/c9rp00241c.

(72) Watts, F. M.; Schmidt-McCormack, J. A.; Wilhelm, C. A.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21* (4), 1148–1172. https://doi.org/10.1039/C9RP00185A.

(73) Grove, N. P.; Cooper, M. M.; Cox, E. L. Does Mechanistic Thinking Improve Student Success in Organic Chemistry? *J. Chem. Educ.* **2012**, *89* (7), 850–853. https://doi.org/10.1021/ed200394d.

(74) Graulich, N.; Hedtrich, S.; Harzenetter, R. Explicit *versus* Implicit Similarity – Exploring Relational Conceptual Understanding in Organic Chemistry. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 924–936. https://doi.org/10.1039/C9RP00054B.

(75) Keiner, L.; Graulich, N. Beyond the Beaker: Students' Use of a Scaffold to Connect Observations with the Particle Level in the Organic Chemistry Laboratory. *Chem. Educ. Res. Pract.* **2021**, *22* (1), 146–163. https://doi.org/10.1039/D0RP00206B.

(76) Rodemer, M.; Eckhard, J.; Graulich, N.; Bernholt, S. Connecting Explanations to Representations: Benefits of Highlighting Techniques in Tutorial Videos on Students' Learning in Organic Chemistry. *International Journal of Science Education* **2021**, *43* (17), 2707–2728. https://doi.org/10.1080/09500693.2021.1985743.

(77) 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.

(78) Darden, L. Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/Backward Chaining. *Philos. Sci.* **2002**, *69* (S3), S354–S365. https://doi.org/10.1086/341858.

(79) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; 2011.

(80) Houchlei, S. K.; Bloch, R. R.; Cooper, M. M. Mechanisms, Models, and Explanations: Analyzing the Mechanistic Paths Students Take to Reach a Product for Familiar and Unfamiliar Organic Reactions. *J. Chem. Educ.* **2021**, *98* (9), 2751–2764. https://doi.org/10.1021/acs.jchemed.1c00099.

(81) Petritis, S. J.; Kelley, C.; Talanquer, V. Exploring the Impact of the Framing of a Laboratory Experiment on the Nature of Student Argumentation. *Chem. Educ. Res. Pract.* **2021**, *22* (1), 105–121. https://doi.org/10.1039/D0RP00268B.

(82) Cooper, M. M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(83) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

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

(85) 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.

(86) Betancourt-Pérez, R.; Rodríguez, J.; Muñoz-Hernández, L. Homing in on the Capabilities That Are Most Predictive of Student Success in the First Semester of Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (3), 635–642. https://doi.org/10.1021/acs.jchemed.9b00568.

(87) Galloway, K. R.; Leung, M. W.; Flynn, A. B. Patterns of Reactions: A Card Sort Task to Investigate Students' Organization of Organic Chemistry Reactions. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 30–52. https://doi.org/10.1039/C8RP00120K.

(88) Galloway, K. R.; Leung, M. W.; Flynn, A. B. A Comparison of How Undergraduates, Graduate Students, and Professors Organize Organic Chemistry Reactions. *J. Chem. Educ.* **2018**, *95* (3), 355–365. https://doi.org/10.1021/acs.jchemed.7b00743.

(89) Lapierre, K. R.; Flynn, A. B. An Online Categorization Task to Investigate Changes in Students' Interpretations of Organic Chemistry Reactions. *J Res Sci Teach* **2020**, *57* (1), 87–111. https://doi.org/10.1002/tea.21586.

(90) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402. https://doi.org/10.1021/ed082p1402.

(91) Grove, N. P.; Cooper, M. M.; Rush, K. M. Decorating with Arrows: Toward the Development of Representational Competence in Organic Chemistry. *J. Chem. Educ.* **2012**, *89* (7), 844–849. https://doi.org/10.1021/ed2003934.

(92) Rushton, G. T.; Hardy, R. C.; Gwaltney, K. P.; Lewis, S. E. Alternative Conceptions of Organic Chemistry Topics among Fourth Year Chemistry Students. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 122–130. https://doi.org/10.1039/B806228P.

(93) Anzovino, M. E.; Lowery Bretz, S. Organic Chemistry Students' Ideas about Nucleophiles and Electrophiles: The Role of Charges and Mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797–810. https://doi.org/10.1039/C5RP00113G.

(94) Bhattacharyya, G.; Harris, M. S. Compromised Structures: Verbal Descriptions of Mechanism Diagrams. *J. Chem. Educ.* **2018**, *95* (3), 366–375. https://doi.org/10.1021/acs.jchemed.7b00157.

(95) 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.

(96) Cartrette, D. P.; Mayo, P. M. Students' Understanding of Acids/Bases in Organic Chemistry Contexts. *Chem. Educ. Res. Pract.* **2011**, *12* (1), 29–39. https://doi.org/10.1039/C1RP90005F.

(97) DeFever, R. S.; Bruce, H.; Bhattacharyya, G. Mental Rolodexing: Senior Chemistry Majors' Understanding of Chemical and Physical Properties. *J. Chem. Educ.* **2015**, *92* (3), 415–426. https://doi.org/10.1021/ed500360g.

(98) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Fragmented Ideas about the Structure and Function of Nucleophiles and Electrophiles: A Concept Map Analysis. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 1019–1029. https://doi.org/10.1039/C6RP00111D.

(99) Kozma, R. The Material Features of Multiple Representations and Their Cognitive and Social Affordances for Science Understanding. *Learning and Instruction* **2003**, *13* (2), 205–226. https://doi.org/10.1016/S0959-4752(02)00021-X.

(100) Schroeder, J. D.; Greenbowe, T. J. Implementing POGIL in the Lecture and the Science Writing Heuristic in the Laboratory—Student Perceptions and Performance in Undergraduate Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 149–156. https://doi.org/10.1039/B806231P.

(101) Stoyanovich, C.; Gandhi, A.; Flynn, A. B. Acid–Base Learning Outcomes for Students in an Introductory Organic Chemistry Course. *J. Chem. Educ.* **2015**, *92* (2), 220–229. https://doi.org/10.1021/ed5003338.

(102) Bretz, S. L.; McClary, L. Students' Understandings of Acid Strength: How Meaningful Is Reliability When Measuring Alternative Conceptions? *J. Chem. Educ.* **2015**, *92* (2), 212–219. https://doi.org/10.1021/ed5005195.

(103) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-to-Learn Assignment in Student Understanding of Organic Acid–Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/C8RP00260F.

(104) Schmid, S.; Youl, D. J.; George, A. V.; Read, J. R. Effectiveness of a Short, Intense Bridging Course for Scaffolding Students Commencing University-Level Study of Chemistry. *International Journal of Science Education* **2012**, *34* (8), 1211–1234. https://doi.org/10.1080/09500693.2012.663116.

(105) 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.

(106) Betancourt-Pérez, 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.

(107) Hermanns, J. The Task Navigator Following the STRAKNAP Concept: Development, Application, and Evaluation of a New Scaffold to Support Nonmajor Chemistry Students While Solving Tasks in Organic Chemistry. *J. Chem. Educ.* **2021**, *98* (4), 1077–1087. https://doi.org/10.1021/acs.jchemed.0c01162.

(108) Stowe, R. L.; Cooper, M. M. Practicing What We Preach: Assessing "Critical Thinking" in Organic Chemistry. *J. Chem. Educ.* **2017**, *94* (12), 1852–1859. https://doi.org/10.1021/acs.jchemed.7b00335.

(109) Graulich, N.; Schween, M. Concept-Oriented Task Design: Making Purposeful Case Comparisons in Organic Chemistry. *J. Chem. Educ.* **2018**, *95* (3), 376–383. https://doi.org/10.1021/acs.jchemed.7b00672.

(110) Graulich, N.; Caspari, I. Designing a Scaffold for Mechanistic Reasoning in Organic Chemistry. *Chemistry Teacher International* **2021**, *3* (1), 19–30. https://doi.org/10.1515/cti-2020-0001.

(111) Gupte, T.; Watts, F. M.; Schmidt-McCormack, J. A.; Zaimi, I.; Gere, A. R.; Shultz, G. V. Students' Meaningful Learning Experiences from Participating in Organic Chemistry Writing-to-Learn Activities. *Chem. Educ. Res. Pract.* **2021**, *22* (2), 396–414. https://doi.org/10.1039/D0RP00266F.

(112) Raker, J. R.; Towns, M. H. Problem Types in Synthetic Organic Chemistry Research: Implications for the Development of Curricular Problems for Second-Year Level Organic Chemistry Instruction. *Chemistry Education Research and Practice* **2012**, *13* (3), 179–185. https://doi.org/10.1039/C2RP90001G.

Chapter 2

Eliciting Students' Reasoning about Acid–Base Reactions With a Mobile Device Application

2.1 Initial remarks

This chapter presents the first of two studies that investigate different instructional supports to elicit organic chemistry students' reasoning about organic reaction mechanisms. As illustrated in Chapter 1, there are numerous studies which investigate students' reasoning with organic reaction mechanisms; however, there are fewer studies which focus on different instructional supports that can elicit students' reasoning. Hence, research focused on how students reason with different instructional supports is needed. Of particular importance is to elucidate whether these instructional supports influence students' reasoning or promote more sophisticated reasoning. This chapter specifically describes an investigation focused on eliciting students' reasoning about acid–base reaction mechanisms as they use a mobile device application designed to support students' mechanistic reasoning.

The study specifically investigated second-semester organic chemistry students' reasoning through think-aloud interviews as students worked through different acid–base mechanism problems using either the mobile device application or paper and pencil. Guided by the models and modelling framework, the analysis of students' responses sought to identify how students' mental models of reaction mechanisms were elicited by the two different modalities. Results from the study indicate how students from both modality groups understand and use underlying concepts related to acid–base mechanisms. Specifically, students were found to focus on explicit rather than implicit features of the reaction mechanisms (e.g., students rarely discussed the reactions in terms of electron movement, an implicit feature). Another key finding was that students often discussed the relevant concepts when approaching the mechanism problems but were not always successful in applying the concepts appropriately (e.g., recognizing the need to use pK_a values, but facing challenges with using pK_a to determine relative acidity). Furthermore, the findings suggested that the different modalities can influence students' problem solving; for example, students working in the application were better able to attempt different steps to solve the problems when their initial attempt was incorrect. This finding is not without drawback, as students using the application did not often reflect on the chemical reasoning for the different mechanistic steps they attempted. Hence, a key implication for practice is to support all students in engaging in strategies to reflect on the chemical reasoning for the mechanistic steps they depict.

This chapter was originally published as a research article in *Chemistry Education Research and Practice*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author on the manuscript, I contributed to methodology, data analysis, and writing (both original draft preparation and review and editing). I shared primary authorship with M.N. Petterson, an undergraduate mentee, who contributed to data collection, analysis, and writing (original draft preparation for sections of the manuscript). E.P. Snyder-White and S.R. Archer, also undergraduates, contributed to data collection and initial stages of data analysis. G.V. Shultz contributed project supervision and writing (review and editing), and S.A. Finkenstaedt-Quinn contributed to funding acquisition, project supervision, conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing).

Original publication and copyright information

2.2 Abstract

An understanding of acid–base reactions is necessary for success in chemistry courses and relevant to careers outside of chemistry, yet research has demonstrated that students often struggle with learning acid–base reaction mechanisms in organic chemistry. One response to this challenge is the development of educational applications to support instruction and learning. The development of these supports also creates an opportunity to probe students' thinking about organic chemistry reaction mechanisms using multiple modalities—i.e., using an app interface or the traditional paper–pencil. This study used think-aloud interviews conducted with undergraduate

students in their first semester of organic chemistry to understand how they worked through two acid–base reactions using either paper–pencil or an app. Analysis of the interviews indicates that students from both groups recognize the steps of acid–base reactions, but do not always apply the underlying concepts, such as assessment of pK_a values or resonance, when determining how a reaction will proceed. The modality seemed to somewhat influence students' thinking, as the app prevented students from making chemically unreasonable mistakes. However, some students relied on the cues it provided, which could potentially be problematic when they are required to respond to assessments that do not provide these cues. Our results suggest that instructors should emphasize the conceptual grounding for the steps that govern acid–base reactions to promote chemical thinking about the relationships between the reaction components and how those influence reaction outcomes, as well as support students to think critically about the chemical information contained within the modalities they are using.

2.3 Introduction

Acid–base chemistry is a fundamental topic in organic chemistry that guides our understanding of chemical reactivity and reaction pathways. Acid–base reactions frequently appear as steps within other reaction mechanisms students learn in introductory organic chemistry.[1] Furthermore, acid–base chemistry was consistently identified as one of the top three most important topics in a study of professors' beliefs about fundamental concepts in organic chemistry.[2] Not only must students have a conceptual understanding of the topic, but they must also be able to apply that conceptual knowledge when reasoning through reaction mechanisms to be successful in organic chemistry.[1,3] Beyond the importance of acid–base chemistry in organic chemistry, an understanding of the topic is also necessary because acid–base reactions commonly appear in other settings such as biochemistry[1,4] and materials chemistry.[5] Reactions mediated by acid–base chemistry are one of the first reaction types covered in the organic chemistry curriculum, and it is within this context that students begin developing the ability to apply conceptual reasoning to reaction mechanisms. Therefore, it is valuable to specifically study how students think about acid–base organic reaction mechanisms.

For research that explores students' thinking about a particular topic, it can be valuable to probe student reasoning using multiple modalities, as the modality may elicit or influence certain thought processes. In particular, with the increase in touch-screen educational software to support students' learning of organic chemistry,[6–13] it is of interest to explore student's thinking about acid–base reactions when working with representations of reaction mechanisms on touch-screen devices as compared to their thinking when working acid–base mechanisms with the conventional paper and pencil. Prior studies have shown how the nature of the task—e.g., the type of problem posed or the way a question is asked—can influence students' reasoning about acids and bases.[14,15] McClary and Talanquer identified that some students use different mental models of acids when performing different tasks related to ranking relative acid strength,[14] and, in a separate study, Cooper et al. demonstrated that the structure of an assessment task influenced the quality of students' reasoning about acid–base reaction mechanisms.[15] While these studies have shown that the way a problem is posed can influence students' thinking about acid–base chemistry concepts, there has been little research into how the modality of a task itself might similarly affect students' thinking due to inherent differences in prompting and structure depiction.

2.3.1 Student understanding of acid–base reaction mechanisms

Organic chemistry typically begins with a re-introduction to the acid–base concepts taught in high school and undergraduate general chemistry courses. Studies have documented common alternative conceptions about acid–base chemistry at these introductory levels,[16] which students might bring into organic chemistry. In addition, the reasoning skills students develop in general chemistry do not necessarily transfer to successful reasoning about acids and bases in organic chemistry.[17,18] For example, Anderson and Bodner identified that while some students can successfully transfer their notions of periodic trends to understand that acids such as HBr and HCl react similarly, their reliance on the location of elements on the periodic table can lead them to classify H_3O^+ as reacting differently than HBr and HCl.[17] Additionally, Cartrette and Mayo identified that students often rely on the Brønsted–Lowry definitions of acids as proton donors and bases as proton acceptors in the context of organic reaction mechanisms, perhaps due to the focus on the Brønsted–Lowry theory during general chemistry instruction.[18] These studies suggest that students are able to transfer knowledge from general to organic chemistry, but they do not always successfully use this knowledge to reason through acid–base reaction mechanisms. This may be exacerbated by the difficulties that students have using pK_a values in the context of organic chemistry reactions.[19] Beyond the lack of successful transfer from general to organic chemistry, the challenges students face with learning acid–base chemistry can persist into graduate school.[20]

Hence, it is necessary to support students' understanding of the different acid–base theories and how to successfully use them for problem solving early in the undergraduate curriculum.[18,21]

Lewis acid–base theory has been found to be particularly important for students' learning of organic reaction mechanisms involving acids and bases because of the theory's focus on electron transfer.[15,22] Corroborating these findings, studies of faculty members' perceptions have identified that understanding Lewis acid–base theory is critical for successful mechanistic reasoning.[23] However, students are often not able to accurately identify Lewis acids and bases, though they are able to correctly identify Brønsted–Lowry acids and bases.[18] Other research has revealed that students have difficulties understanding, applying, and describing reactions in terms of the electronics inherent to Lewis acid–base theory.[18,24] Furthermore, students have many mental models of acids and bases and they often struggle to switch between models.[14] In particular, when considering acid strength, students tend to focus primarily on surface features related to the Arrhenius and Brønsted–Lowry acid–base theories—such as the presence of dissociable protons—rather than the implicit electronics of Lewis acid–base theory, and only invoke the Lewis theory in conjunction with mental models related to the other two theories.[14,22]

Taken together, the prior research on students' conceptions of acids and bases suggests that students struggle to apply Lewis acid–base theory in comparison to other theories. This is potentially troubling in the context of organic reaction mechanisms, as both Lewis acid–base theory and organic reaction mechanisms involve explaining reactions based on the movement or transfer of electron pairs. The focus on electron transfer in the Lewis acid–base theory leads into an understanding of mechanisms more generally, as the Lewis theory allows for an electronic explanation of how proton transfers occur.[15] Electronic explanations of mechanisms are necessary for mechanistic reasoning in organic chemistry,[23] and it is therefore valuable to understand if and how students are using the Lewis theory to think about acid–base reaction mechanisms. This foundation is particularly important because conceptual understanding of acid–base reaction mechanisms lends itself to better understanding of other reaction mechanisms, such as nucleophilic additions.[1,15,18,21]

2.3.2 Conventional **versus** *touch-screen interfaces in organic chemistry*

Line-angle structures are the conventional method for presenting organic molecules. Students often work mechanism problems by drawing arrows from nucleophilic to electrophilic sites represented in the line-angle structures. In addition to line-angle structures, interfaces on touch-screen devices also exist that allow students to construct and manipulate organic structures.[7,9] One such application, "OrganicPad," allows students to construct Lewis structures and place arrows to illustrate one-step reaction mechanisms.[7] After drawing Lewis structures, students can direct the application to check for possible mistakes or convert their two-dimensional representations into three dimensions.[7] "OrganicPad" has been used in research settings to identify challenges students face with drawing Lewis structures[8] and with drawing static reaction mechanisms.[6] A similar application, "Molecules," allows users to manipulate two-dimensional projections of three-dimensional ball-and-stick and space-filling representations of organic structures using a touch screen.[9] This application has been shown to improve students' representational competence skills.[11] While these applications have been shown to support students' learning of organic representations, there has not been research focused on applications that specifically target the process of organic reaction mechanisms.

A recently-developed app, "Mechanisms," can act as a tool for students studying organic reaction mechanisms.[12,25] It encompasses a comprehensive range of mechanisms including acid–base, addition, substitution, elimination, and electrophilic aromatic substitution reactions. The app models atoms, bonds, and electrons in a way that allows the user to dynamically manipulate chemical structures over the course of a mechanism. This interactive interface allows users to tap on carbon atoms to reveal implicit hydrogen atoms and to tap on heteroatoms or carbanions to reveal non-bonding electron pairs. Students are able to form bonds by dragging electron pairs from bonds or atoms to another atom, and the app shows users the chemical feasibility of the electron movements in real-time by either allowing the new bonds to form or by rejecting the electron movements and returning the electrons to their source. The app also provides students with guidance towards correct product formation through task cards, goals, and hints, which give information about the reaction. Since the app offers a different modality for students to work through reaction mechanisms—a modality which inherently presents reactions differently and provides additional prompting compared to the traditional paper–pencil modality—it is valuable to explore students' thinking when using this modality as it may elicit a greater range or different types of conceptions. The app's interactive interface could be of particular interest in light of the Bongers et al. finding that students developed more dynamic mental models of reaction mechanisms following a learning activity that incorporated animated, as opposed to static, representations of a reaction mechanism.[26] As such, the present study focuses on exploring students' thinking—the chemical features and concepts they consider—when working through acid–base organic reaction mechanisms using either the "Mechanisms" app interface or the traditional paper–pencil.

2.4 Theoretical framework

This research is guided by the models and modelling framework originally derived from the Lesh et al. formulation of mental models[27] and adapted by Briggs and Bodner[28] and Briggs.[29] This framework separates mental models into five components: (1) referents, (2) relationships, (3) rules/syntax, (4) operations, and (5) results.[28,29] Referents are specific representations or symbols, such as atoms or molecules. Relationships are how referents relate to one another, either within molecules (e.g., atoms within a molecule relate to one another through bonds) or between molecules (e.g., the relative acidity or basicity of two molecules). The relationships are dictated by rules and syntax, where rules are defined as concepts and syntax as how rules are utilized in a task.[28,29] In our context, an example of a rule is the concept that bases donate electron pairs to acids, and syntax would be knowing to consider the relative acidity and basicity of sites on a molecule—using other concepts such as pK_a values and resonance—when determining which atom will donate or accept electron pairs. Operations are how referents are manipulated by applying relationships and rules to produce new representations. For example, an operation would be the action of applying the rules and syntax related to acidity and basicity to protonate the base present in the reaction. Lastly, results are the outcomes of the operation which can be used as a source of new knowledge that may inform future steps (e.g., the result of a reaction intermediate with a new set of properties that can be used to guide decisions about the next step of a reaction). Operations are unique in that they are a dynamic component whereas the other components are static.

The models and modelling framework provides a lens for examining the chemical features that students consider and apply when working through organic reaction mechanisms. The ability to identify the key referents and the relationships between them and then apply the appropriate rules and syntax allows students to proceed through a reaction mechanism as a series of chemically correct and favored operations. With each new result, students have to take into account how the components may have changed to determine the next operation to perform and to know when they have reached the final result or product. Not only may there be variation across reactions in how students use the components of mental models, but the way information is presented may also elicit different modes of thinking or influence how students utilize the components of mental models. For example, students may engage differently with the representation of referents in the modalities explored herein, as lone electron pairs that are drawn explicitly on paper are hidden in the app unless students tap on atoms to reveal them. Additionally, the two modalities contain specific prompts that are inherent to them which may influence which components of the framework students use as well as how they use them. For example, in the app, the results of some incorrect operations are either not allowed or lead to hints that act as cues to the relationships, rules, and syntax important to the reaction. Thus, probing and analyzing student thinking *via* multiple modalities, and situating this analysis in the models and modelling framework, provides a better understanding about how students think about reaction mechanisms.

2.5 Research questions

This study investigated how first semester organic chemistry students reason through acid–base reaction mechanisms when completing tasks *via* different modalities. To do this, we had students think aloud while working through two acid–base reaction mechanisms. Students were assigned to one of two groups, where one group worked through the reactions on paper and the other group worked reactions with the "Mechanisms" app. The following research questions guided our investigation:

1. How are students in organic chemistry reasoning when using either a touch-screen application or the traditional paper–pencil method when working acid–base reaction mechanisms?
2. What components of mental models do students focus on when reasoning through acid–base reaction mechanisms?

2.6 Methods

2.6.1 Context and participants

The study was conducted at a large, Midwestern research university. Students were recruited using a mix of purposeful and convenience sampling[30] across three semesters from the first of a two-course, lecture-based introductory organic chemistry sequence. Brønsted–Lowry acid–base reactions are the first reaction types covered in the course, following a review of relevant general chemistry content and an introduction to resonance, VSEPR and MO theory, and the curved arrow notation. Students were recruited prior to the first exam, which also covered electrophilic addition reactions. Students were expected to be able to identify strong *versus* weak acids and bases, identify the most acidic proton or basic atom in a structure, use the pK_a table to determine approximate pK_a values and to identify whether structures are protonated or deprotonated given the pH of a solution, and draw mechanisms for acid–base reactions. During the first semester of data collection, students were recruited using a list provided by the instructor of the course which contained the names of the students from the top and bottom pools of scores from the first exam. This allowed for purposeful selection so that participants would have a range of abilities and conceptions. During the second and third semester of data collection, students were recruited by a course announcement for convenience sampling to increase the number of participants in the study. During the recruitment process, students were told that participating in the study would provide them with practice on organic chemistry mechanisms and, following working through the reactions, that they would be able to ask the interviewer any organic chemistry related questions they had. No additional incentives were provided. In total, thirteen students were recruited to participate in think-aloud interviews. Six of the students worked through the reaction mechanisms using the conventional paper–pencil method, denoted as paper–pencil students, and seven worked through mechanisms using the "Mechanisms" app, denoted as app students. Students were randomly assigned pseudonyms that are not representative of their ethnicity, gender, or other identities (Table 1.1). The research team received Institutional Review Board approval (HUM00156602) for the data collection and analysis in this study. Students consented to be part of the study at the beginning of the think-aloud interviews.

Table 2.1 Student participants by think-aloud interview group type

Reaction modality groups	**Participants**
Paper–pencil	Ana, Aurora, Daisy, Francis, Mary, Perdita
App	Angela, Belle, Flynn, Jasmine, Pepper, Peter, Tiana

2.6.2 Reaction selection

We selected reactions from the app based on the reactions covered in the course. The app presents students with the reactants (2.12.1 Appendix 1, Figure 2.3) but does not show the target products; however, each puzzle starts with a task card that shows mechanistic arrows indicating moves students will have to make or intermediates of the reaction. Additionally, the app may present students with hints and goals during the puzzles to direct students toward the desired products (2.12.1 Appendix 1, Figure 2.3). To mirror the level of information that students received from the app, we depicted the reactions for the paper–pencil students by presenting the line-angle representation of the organic reactants and the molecular formula of the major product, with the additional reagents depicted above the reaction arrow. To assess the content validity of the chosen reactions, we discussed them with three instructors for the course, one who was teaching the course during the first semester of data collection and two who had previously taught the course at the study institution. They felt the chosen reactions were similar to those students would be expected to solve and were at an appropriate difficulty level. Additionally, input from expert organic chemistry instructors guided the translation of presenting the problems within the app to the presentation on paper, to ensure students' responses were reflective of how students would be thinking when working with these different modalities in authentic settings (e.g., while studying for an exam). We discussed the presentation with one instructor, made adjustments, and confirmed with the other instructors that the approach would not cause students undue difficulty in interpreting the questions and that they were similar in terms of the initial information provided by the app. For example, the molecular formulas of the major products, but not the minor products, were provided to the paper–pencil students in an effort to mitigate the advantage tendered to the app students *via* the provided hints and goals. Additionally, the reactions were unbalanced due to similar reasoning. The instructors verified that students should be familiar with reactions presented in this form, with both the lack of minor products and balancing mimicking how reactions are sometimes presented in organic chemistry lectures and textbooks. The final selected reactions are depicted in Figure 2.1 and Figure 2.2.

2.6.3 Think-aloud interviews

Interviews followed a think-aloud procedure, where students were prompted to verbalize their thinking as they worked through the series of reactions.[31,32] Each think-aloud interview consisted of students working through four organic chemistry reaction mechanisms, either on paper or using the app. Results from the two acid–base reaction mechanisms are presented herein. At the beginning of the interview, students did a practice think-aloud to acclimate them to verbalizing their thoughts. During the think-aloud interviews, interviewers used probes such as "Why did you make that move?" or "What are you thinking about right now?" to prompt students to explain their reasoning. Additionally, all students were provided with the pK_a table used in their organic chemistry course for reference as, in this institutional context, it is a resource they receive at the beginning of the semester and during course assessments. The pK_a values from the table relevant to the two reactions discussed herein are presented in 2.12.2 Appendix 2, Figure 2.4. For each student, order of the reactions was randomized. All of the interviews were video and audio recorded.

In the paper–pencil think-aloud interviews, students used a Livescribe™ pen and notebook, which recorded their writing in real time. Data collected with the Livescribe™ supplemented the audio and visual data. Prior to each interview, the interviewer wrote the reactions on separate pages in the Livescribe™ notebook in random order. Students were prompted to write all their work in the notebook and could use additional pages if necessary. To align how the reactions were presented to the app and paper–pencil students, the paper–pencil students were told the type of reaction they were doing prior to starting each reaction, as the reaction type was given in the task card presented by the app. Additionally, paper–pencil students were asked at the end of the reaction whether there were any resonance structures relevant to the reaction, as the app prompted students to show all resonance structures. We did not provide explicit cues to students to parallel the other prompts that were provided by the app (e.g., hints).

Interviews with the app students were conducted similarly to paper–pencil interviews with the addition that students were given an abbreviated version of the tutorial provided by the app before starting the think-aloud interview. The tutorial was adapted by one member of the research team (ESW) and refined by independently piloting it with two other members of the research team (SFQ and MP) who had not yet used the app. The tutorial instructed students on how to reveal implicit lone pairs and hydrogen atoms, how to create and break bonds, and how to move and rotate molecules. This ensured that unfamiliarity with the app's functions did not inhibit students' abilities to work through the reactions. Two of the app students had used the app previously and the remaining app students did not exhibit undue difficulty. An occasional difficulty students encountered when using the interface was getting the app to register their intended movements of electron pairs. When a student made a correct move that the app did not register as such, the interviewer suggested they try again as the difficulty was not related to the student's thinking about the chemistry.

2.6.4 Development and application of the coding scheme

The coding scheme was developed through open coding and constant comparison of the think-aloud interviews.[33] Four of the researchers (SFQ, MP, ESW, SA) reviewed the transcripts and audio/visual data produced from the think-aloud interviews, noting observations related to students' thinking and identifying initial codes. The research team discussed the codes and grouped them into parent codes of chemical considerations, reaction step, participant usage, justification, student actions, and app-specific. Two of the four researchers (SFQ and MP) then finalized the coding scheme and trained a fifth member of the research team (FW) to use the coding scheme. The coding scheme is presented in 2.12.3 Appendix 3, Table 2.3.

To establish what sections of each transcript should be coded, all transcripts were divided into units of analysis corresponding to thinking stages, where students verbalized their ideas about steps in the reaction, and action/operation stages, where students performed the electron movements to break and form bonds. The two members of the research team who finalized the coding scheme (SFQ and MP) identified and agreed upon the units of analysis for all transcripts before coding. One of those researchers (MP) and the trained fifth member (FW), who was not involved in the development of the coding scheme, then independently coded both reactions from four participants (30% of the data), met to clarify the coding definitions, and came to a consensus on the application of the coding scheme for these reactions. Afterwards, the same two researchers (MP and FW) independently coded both reactions from the remaining nine participants (70% of the data). During this process, the researchers met to discuss the application of the coding scheme, assess agreement using the fuzzy kappa statistic,[34] and come to a consensus for coding. The initial fuzzy kappa value for the 70% of the data coded after clarifying the coding scheme was 0.82, within the range indicating near-perfect agreement.[35] Furthermore, as consensus was reached for each transcript, the researchers overcame initial coding disagreements to achieve complete agreement.

2.7 Results

The results are drawn from the qualitative analysis of students' think-aloud interviews in which they attempted to produce the mechanisms for two acid–base reactions using one of the two modalities. This analysis was guided by the models and modelling framework, and thus we refer to atoms and molecules as referents, the concepts students draw upon as rules, and the way students apply concepts as syntax. By examining the rules/concepts students referred to and the syntax with which they applied these rules, we are able to identify the reasoning students exhibited when considering the mechanisms. Analyzing the interviews through the lens of the models and modelling framework additionally allows us to begin differentiating whether students' difficulties arise from their conceptual knowledge or their ability to apply that knowledge. Furthermore, we examine how the two modalities, and the prompts inherent to each, may influence student reasoning. We first present students' responses when producing a mechanism for the deprotonation of a 1,3-dicarbonyl, followed by students' responses when producing a mechanism for the protonation of imidazole.

2.7.1 Deprotonation of a 1,3-dicarbonyl by a strong base

In this reaction, students first needed to assign the roles each molecule would play (i.e., acid or base) by determining the relationship between the referents. Then, considering the rules and syntax associated with acid–base chemistry, they needed to identify the most acidic site for deprotonation on the dicarbonyl (Figure 2.1). The pK_a table all students were given included, among pK_a values for other structures, a dicarbonyl similar to that in the reaction and the pK_a value for water which they could use to identify relative acidity and basicity should they need it as a resource (2.12.2 Appendix 2, Figure 2.4A). Following their decisions about acidity and basicity, students could then perform the associated operations, where the result should lead to a consideration of resonance stabilization of the product. The students in each group tended to approach each step of the mechanism using distinct reasoning, potentially due to differences in prompting by the modalities, and thus they will be discussed separately.

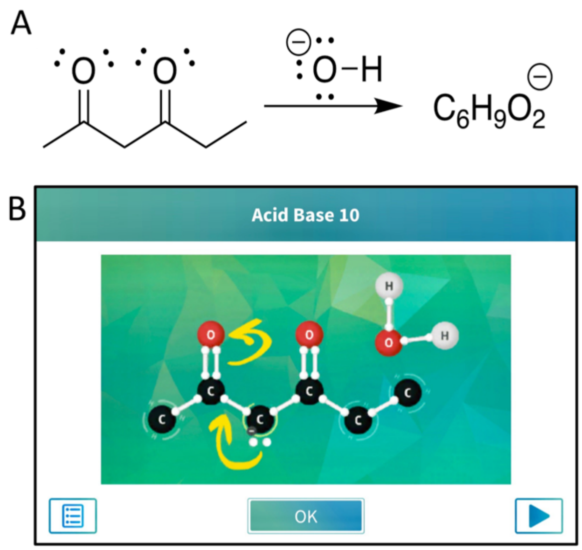


Figure 2.1 Reaction schemes for the deprotonation of a 1,3-dicarbonyl by a strong base as presented during the think-aloud interviews to (A) paper-pencil students and (B) app students in the task card prior to beginning the reaction.

Most paper–pencil students started the reaction by attempting to determine the acid–base relationships between the molecules in the reaction. One student, Ana, used an atom-counting strategy to determine that the dicarbonyl compound would lose a proton and then identified that hydroxide would remove the proton. All other paper–pencil students who completed the reaction used the rules of the pK_a table to determine the acid–base relationship between the molecules, where only one student, Francis, first correctly identified the acid and the base using chemical thinking and then confirmed their decision with the pK_a table. Of the students who went directly to the pK_a table to identify each species, Mary correctly identified the role of each species. Daisy and Aurora, however, had some difficulties identifying the acid–base relationship and exhibited an incomplete knowledge of the syntax for using pK_a values in doing so. Aurora incorrectly identified the dicarbonyl as a base and hydroxide as an acid when first looking at the structures, and then turned to the pK_a table to identify the relevant pK_a values. Aurora then started to doubt their original assignment of acid and base, but resorted to using the formula of the major product to determine that the dicarbonyl was losing a proton and must be the acid in the reaction rather than basing their reassignment on the pK_a values. Daisy correctly identified the acid and base using values from the pK_a table, but then revealed incorrect understanding of the underlying concepts when considering how the species would react:

"So, since it's an acid, that means it gets protonated. So, this bond between the OH would break. And then the lone pairs go on the oxygen... And this hydrogen would now be added to one of these. One of the oxygens with the lone pair."

After completing these steps, Daisy counted atoms and identified a discrepancy between the product they had drawn and the given condensed formula, but did not know how to address this discrepancy and stopped working on the reaction. While for Aurora the pK_a values cued a discrepancy with their original assignment of acid and base, Daisy was not able to move from the pK_a values to what they indicated about which species was donating or accepting a proton.

One paper–pencil student, Perdita, did not attempt the problem, initially approaching the reaction similarly to Aurora by first considering the carbonyl oxygen atom as a base and then using an atom-counting strategy. However, as side-products were not shown and the presented reaction was not balanced, Perdita did not know how to account for the apparent loss of an oxygen atom:

"Well, I guess I'm confused in general, because there's three oxygens over here, and then over here there's only two. So I'm like, where does this third oxygen go? Which I'm confused about. So. . . I don't know, an oxygen just vanishes."

Although Perdita did not complete the reaction, they did initially attempt to identify the acid–base relationship. Perdita recognized their initial assignment of acid and base to be incorrect, but then did not attempt the reaction further after not knowing how to navigate the unbalanced reaction. Perdita's difficulty with how the paper–pencil representation was presented is important to note, as instructors and textbooks do not always provide students with balanced reactions.

The app students were more varied in how they began the reaction. Few students began by attempting to determine the acid–base relationship and only one student, Belle, correctly identified the acid and the base, noting the charge on the hydroxide and using the pK_a table to guide their thinking. Tiana immediately looked at the reacting species and the pK_a table and incorrectly identified the hydroxide hydrogen atom as the most acidic proton. However, after attempting an electron movement the app did not allow, Tiana examined the task card and immediately realized the appropriate mechanistic step. Angela also struggled to identify the acid and base, recognizing both the hydroxide and the carbonyl oxygen atoms as having lone electron pairs and capable of being protonated. Notably, Angela did not use the pK_a table to guide their thinking, instead attempting to protonate one of the carbonyls—a move the app would not allow—before turning to the goals within the app to help guide their thinking. The remaining app students immediately relied on the task card that was presented to them at the beginning of the reaction to guide their first steps, effectively skipping the step of identifying the relationship between the molecules as the task card indicates which molecule gains and which loses the proton that is transferred during the reaction (Figure 2.1B).

For the paper–pencil students who identified the acid and the base in the reaction, the next step was to use the rules and syntax of acid–base reactions to determine the operation of which proton would be removed from the dicarbonyl compound. They primarily used the pK_a table, with some also considering the rules and syntax associated with resonance to make this decision. Mary and Francis used the pK_a table to identify the appropriate proton to be removed. Mary commented on the difference in pK_a values between the acid and the conjugate acid of the base to confirm their choice and, while they did deprotonate at the correct site, did not consider which protons adjacent to the carbonyls were the most acidic. Francis considered other protons that could be removed from the dicarbonyl, but justified that one of the protons in between the two carbonyls would be removed because they recognized that deprotonation between the two carbonyls would result in a product that could be stabilized by resonance. Aurora and Ana, also paper–pencil students, recognized the need to consider which of the protons adjacent to the carbonyls would be removed and considered resonance to guide the decisions they made. However, both neglected to consider the protons in between the two carbonyls. Aurora started to consider the correct protons following probing about why they had considered the protons they initially focused on. After this probing, they identified the oxygen atoms in the carbonyls as allowing the potential for resonance stabilization in the deprotonated product, and then used the pK_a table to confirm which were the most acidic, ultimately deprotonating the correct carbon atom:

> "Yeah, I guess it could also come off here, that might actually be more stable. I don't know if there [is] an exact pK_a —oh wait, this is kind of. . . this is 9.2, this is the one for the hydrogen right there, so that would probably be it because that's more stable because there's more resonance coming from both these O's."

Despite also consulting the pK_a table and considering the possibility for resonance structures in the deprotonated product, Ana ultimately did not use the appropriate syntax for these concepts and chose to deprotonate the incorrect carbon atom.

The majority of the app students who relied on the task or goal cards did not consider which proton to remove when performing their first operation. The task card showed an intermediate step rather than the first step of the reaction, presenting a molecule of water next to the dicarbonyl with a lone electron pair and negative charge at the central carbon atom (Figure 2.1B). Jasmine, Pepper, Angela, and Tiana used the task card to guide their reasoning to deprotonate at the appropriate location without vocalizing any chemical thinking about the rules or syntax of acid–base reactions. In addition to using the task card to guide their initial steps, Flynn and Peter used some chemical thinking to identify the most acidic proton. Both recognized from the task card that the reaction used hydroxide to form water, after which Flynn used the pK_a table to correctly identify the most acidic proton while Peter identified that forming a carbanion adjacent to one of the carbonyls would result in a lone pair that could be delocalized. However, Peter made the same mistake as Aurora and Ana in the paper–pencil group and initially tried to remove a proton that would result in a structure with less resonance stabilization. Since the app did not allow Peter to make this move, Peter then consulted the pK_a table and used the information provided to identify which proton to remove.

After the operation of deprotonation, the final step of the reaction was to use the rules and syntax affiliated with resonance to identify the two primary resonance contributors for the product. All three of the paper–pencil students who deprotonated at the appropriate carbon atom on the dicarbonyl were able to complete this task without difficulty, and most described their reasoning in terms of electronegativity. Following deprotonation, Francis and Mary both drew one of the resonance contributors to show stabilization of the negative charge on the carbon atom. Aurora provided similar reasoning following a post-reaction interview question about the potential for resonance structures. In their discussions, both Francis and Aurora expressed incorrect understanding about resonance. Aurora considered drawing both resonance contributors, but felt that one structure was more stable than the other, conflating stability with degree of contribution to the resonance hybrid. When considering the possibility of the second resonance contributor with the negative charge on an oxygen atom, Francis revealed a misconception regarding resonance structures: "Oh you would have a mixture, because you would always have a mixture. . . like all three of these could still exist in solution."

Only one app student, Belle, showed the resonance structures without being prompted by the app. Belle realized that the carbanion produced was not very stable and was able to depict the two resonance contributors where the negative charge was on one of the carbonyl oxygen atoms which stabilized the structure. The remaining app students required prompting from either the task or goal cards before showing both resonance structures. Only Jasmine and Tiana explicitly expressed that the presence of resonance contributors would stabilize the product, as it places a partial negative charge on the more electronegative oxygen atom. Angela had some difficulties showing the resonance structures, struggling to identify the correct place to start the movement of electrons, first using the lone pairs on the carbonyl oxygen atom before realizing that they needed to start drawing the resonance structures from the lone pair on the negatively charged carbon atom.

In all, students exhibited differences in approach to this reaction depending on whether they were working with the app or with paper-and-pencil. The paper–pencil students tended to begin by trying to identify the acid–base relationship, while app students often skipped this step due to the intermediate structure being provided in the task card for the reaction. Similarly, students from the app group were able to determine the site of deprotonation using the app's guidance, a task which proved challenging for many paper–pencil students. Students across both groups tended to use the rules and syntax of resonance to identify the resonance structures for the product without difficulty, though some did exhibit problematic thinking.

2.7.2 Protonation of imidazole by a strong acid

In the strong acid protonation of imidazole (Figure 2.2) students had to identify the most basic nitrogen atom in the ring by applying the rules and syntax associated with acid–base chemistry and resonance. The key to this reaction was for students to recognize that, after the first operation of protonation, the positive charge on one of the nitrogen atoms would be stabilized through resonance whereas the other would not, indicating the preferred product. The pK_a table that students received had two potential structures they could identify as structurally similar to the two nitrogen atoms in the ring and use to guide their thinking (2.12.2 Appendix 2, Figure 2.4B). Unlike in the dicarbonyl reaction mechanism, where the paper–pencil and app students appeared to make relatively distinct moves, the students approached the imidazole reaction more similarly across the groups and thus will be discussed together.

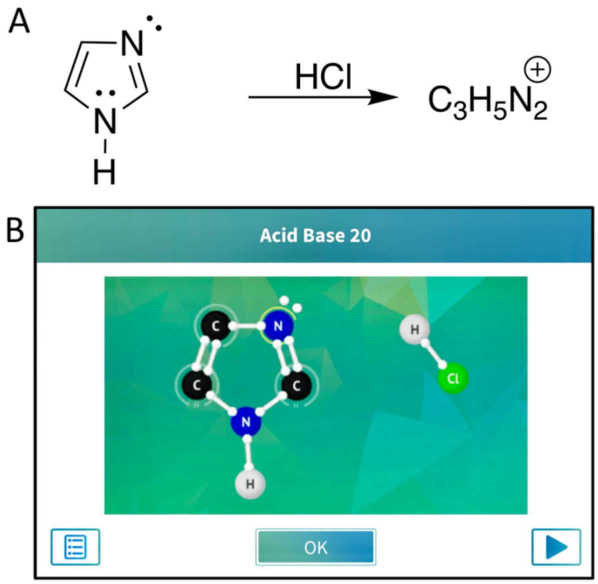


Figure 2.2 Reaction schemes for the protonation of imidazole by a strong acid as presented during the think-aloud interviews to (A) paper-pencil students and (B) app students in the task card prior to beginning the reaction.

Most students from both groups began this reaction by recognizing HCl as a strong acid and using their knowledge of the acid–base relationship to identify that one of the nitrogen atoms in the imidazole ring would be protonated. While most students did not provide a thorough explanation for why a particular nitrogen atom would be protonated, a few students cited reasons for why nitrogen rather than one of the carbon atoms would be protonated. Tiana considered the relationships between the two types of atoms by comparing their basicity, mentioning that nitrogen is more basic than carbon. Aurora reasoned that carbon should not receive a charge and Jasmine identified that the carbon atoms were closed shell, leading both to conclude that a carbon would not be protonated. This indicates that students have some ability to correctly identify basic sites, but it is unclear whether this is from recognizing atoms they are familiar with from other acid–base reactions or if they are actually thinking about chemical properties.

The majority of students generally struggled with the rules and syntax when determining which nitrogen atom to protonate during the first operation. Overall, students in both groups showed a heavy reliance on the pK_a table to determine the correct site for protonation (2.12.2 Appendix 2, Figure 2.4B). Aurora, Daisy, Belle, and Flynn, two students from each group, each only identified one relevant pK_a value on the table and chose to protonate at the corresponding nitrogen atom in imidazole. The thinking behind this was verbalized by Aurora and Flynn, who reasoned that the relevant pK_a values are either given in the table or provided in the reaction. Aurora said:

> "Yeah, I mean, I feel like a lot of times if they don't have it on the pK_a table and it's really important then they give you that value in the question, since the value's not in the question it makes me think that maybe it's not it. Which probably isn't a very good answer, but in a test situation that's probably would I would do."

While three of the four identified the correct nitrogen atom to protonate and were able to proceed, Flynn identified the conjugate acids of ammonia and methylamine in the pK_a table and determined that the pK_a of the secondary amine in the ring would fall between the affiliated pK_a values. Flynn tried to protonate at that nitrogen atom but was prevented by the app. Mary did identify two nitrogen-containing structures in the pK_a table; however, the more basic structure they identified was not a good approximation for the protonated nitrogen atom in imidazole that they related it to. This led Mary to protonate the incorrect nitrogen atom and form the incorrect product. Both Angela and Pepper, app students, did not rely on the pK_a table or initially exhibit chemical reasoning. Angela chose the incorrect nitrogen atom without verbalizing their reasoning before being cued by the app to consider which nitrogen atom was the most basic; Pepper based their decision on the task card for the reaction which showed the lone pairs on the most basic nitrogen atom (Figure 2.2B). After a probing question by the interviewer, both students discussed how they thought the nitrogen atom they did not protonate would be less basic because it already had a hydrogen atom attached.

The remaining students, three from each group, thought about how the rules and syntax of resonance would impact which nitrogen atom was favored for protonation. However, only Francis and Ana, paper–pencil students, recognized that for this reaction they should be considering the potential for resonance in the products and drew potential resonance contributors. Ana said: "So now I have to see which of these structures is better, or which N can hold the positive better." Peter, Tiana, Jasmine, and Perdita all focused on resonance stabilization of the reactant rather than the possible products, incorrectly applying the syntax of resonance structures and ultimately selecting the incorrect nitrogen atom to protonate. Of the four, only Perdita was a paper–pencil student and proceeded to form the incorrect product. Peter, Tiana, and Jasmine received a hint from the app that they should use the most basic lone pair and show delocalization of the positive charge through resonance. While this did not lead them to reason through why their original thinking was incorrect, they did subsequently protonate the correct nitrogen atom. The focus on resonance stabilization of the reactant indicates a gap in students' understanding of how to appropriately apply the syntax of resonance when considering acid–base reaction mechanisms.

Following the operation of protonating the nitrogen atom, students were prompted to draw resonance structures for the resulting product molecule either by the app or, in the case of the paper–pencil students, as part of post-reaction interview questions. Students from both groups had difficulty with this. All of the students except Mary recognized that the product would be stabilized by the presence of resonance contributors, but most students had some difficulty identifying what source of electrons to use when performing the operation to depict the resonance structures. All of the app students, except Flynn, and three of the paper–pencil students tried to start depicting resonance structures from one of the carbon–carbon double bonds in the imidazole rather than using the available lone pairs on the nitrogen atom. Two of the remaining paper–pencil students, Aurora and Mary, did not draw resonance structures; for Mary, this was because they had drawn an incorrect product that did not have the potential for resonance. Francis, the last paper–pencil student, did use the lone pair electrons on the neutral nitrogen atom to start their resonance structures. For the app students, the focus on the double bonds may have been exacerbated by the fact that the lone pairs are not automatically visible in the app and students first had to select the nitrogen atom to reveal them. This is especially interesting as all the paper–pencil students had drawn in the lone pairs present in their final products. This could indicate a focus on the explicit features, such as double bonds, present in the referents and that the app students had difficulty in readily identifying the implicit lone pair electrons on the neutral nitrogen atom.

Overall, this reaction was potentially more difficult for students. Many struggled to apply the rules and syntax of acid–base chemistry and resonance, which led them to protonate the incorrect nitrogen atom during the first operation, or exhibited minimal reasoning when they chose the correct one. The potential for resonance in the product also caused difficulties, where some students recognized the rules of resonance stabilization but they struggled to apply the syntax in predicting the reaction outcome and when depicting the resonance structures of the product.

2.8 Discussion

This research used two modalities, paper–pencil and app, to elicit student reasoning about acid–base organic chemistry reactions. By describing the results through the lens of the models and modelling framework we can characterize what chemical features and concepts students identified as important for reaction progress and how those informed the mechanistic steps they made. This framework also allows for an initial understanding of whether the different representations, or modalities, resulted in different use of the models, which is worth investigating further. We present differences and similarities between students' responses when using the two modalities, and we emphasize that these differences may also stem from differences between the modalities in both how the reactions are presented and how different levels of feedback or prompting are provided. Generally, the students using the app and paper–pencil modalities exhibited commonalities in the chemical features they focused on but appeared to have differences in their approaches, in particular for the dicarbonyl reaction. This may be due to the fact that the presentation of the reaction, which is inherently connected to the modality, may have guided students' thinking. Beyond differences in how the reactions are presented between modalities, differences in students' thinking may also stem from the level of feedback provided within the app compared to the minimal level of feedback when working with paper and pencil. Hence, we consider how the modalities as a whole influence students' reasoning. The common problematic thinking that students demonstrated across both groups and for both reactions are summarized in Table 2.2.

Table 2.2 Common student difficulties across modalities

Problematic student thinking	**Problem**	**Level(s) in the models and modelling framework**
When identifying acids and bases, limiting considerations to surface features and/or Brønsted–Lowry definitions	Dicarbonyl	Relationship, rules and syntax
Not considering the relative acidity of hydrogen atoms	Dicarbonyl	Syntax
Identifying resonance structures as a mixture rather than contributing to a resonance hybrid	Dicarbonyl	Rules
Overreliance on the pK_a table	Imidazole	Rules and syntax
Inability to generalize from the structures provided in the pK_a table	Imidazole	Rules
Focusing on resonance in the reactant rather than the potential product	Imidazole	Syntax
Difficulty drawing resonance structures	Imidazole	Syntax, operations

2.8.1 Students generally focused on explicit, rather than implicit, referents and relationships

Generally, students discussed the reactions in terms of the molecules and atoms involved, using minimal language to describe the breaking and forming of bonds or the movement of electrons. The lack of students using language to describe electron movement to break and form bonds is in contrast to other studies,[36,37] though it does support the finding that students often devalue the physical meaning behind the electron-pushing formalism.[38] When students did talk about electrons, they were often referring to lone pairs available to participate in reaction steps. This supports previous research that indicates students focus on the explicit referents in reactions rather than more implicit features.[36,39–43]

When solving either acid–base reaction, students generally began their thinking by identifying the relationships between referents in the reaction by assessing relative acidity and basicity. Students had more difficulties identifying the acid and base for the dicarbonyl reaction. This could be due to the fact that the acid in the reaction—the dicarbonyl—did not have explicit hydrogen atoms to signal students toward thinking about its relative acidity when combined with hydroxide in the reaction. Similarly, although the hydroxide presented to students in the dicarbonyl reaction had a negative charge, many students did not immediately recognize it as a base and some students mislabeled it as an acid. That students mislabeled hydroxide as an acid is similar to Anderson and Bodner's finding that students incorrectly transfer knowledge of periodic trends when identifying acidic species.[17] Furthermore, the difficulties students had identifying the base despite the presence of a negative charge is suggestive that students were not considering the ability of the reactant to donate electron pairs, aligning with the finding of Cartrette and Mayo that students focus on the Brønsted–Lowry definitions of acids as proton donors and bases as proton acceptors.[18]

Similarly, for the imidazole reaction, students tended to determine the acid–base relationship using surface features of the molecules given: the presence of HCl and of nitrogen atoms in the ring. Hydrochloric acid is likely one of the first strong acids that students learn in general chemistry, and many students immediately recognized it as an acid. Similarly, many students explained that they knew nitrogen atoms in molecules tended to act as basic sites. Students' thinking appeared to be guided by the surface features of these molecules, and as a result they tended to not explicitly consider any specific theory of acids and bases. This is similar to prior findings in the literature in which students were found to make decisions about organic reaction mechanisms by focusing on the surface features of the reactants rather than the chemical information communicated by the structure.[14,40] Students in particular were not considering the Lewis acid–base theory, focusing on the atoms and molecules themselves rather than the ability of reactive species to accept or donate electrons, a finding similar to those in prior research.[18,22,24]

The different levels of ease with which students were able to determine the acid and base between the two reactions may explain why the groups of students were similar in their responses to the imidazole reaction but dissimilar in their responses to the dicarbonyl reaction. Specifically, most students automatically identified HCl as the acid in the imidazole reaction but they had difficulty assigning acid–base character in the dicarbonyl reaction and so relied more heavily on the supports available to them. For the paper–pencil students this was the pK_a table, but the app students were also able to rely on the modality itself as a source of information.

2.8.2 Students recognized the rules related to the reactions, but could not always successfully apply the affiliated syntax

For both reactions, students generally recognized the rules, or pertinent concepts, for the reaction—knowledge of pK_a values, resonance, and that the reaction would involve one species deprotonating another. However, students' recognition of the syntax—of the need to use knowledge of pK_a values and resonance to make a decision about reactivity—differed between reactions. It is important for students to know both the rules and the syntax affiliated with acid–base reactions, as acid–base concepts are frequently utilized in more complex organic chemistry reactions.[1] For the dicarbonyl mechanism, most paper–pencil students knew to use the pK_a table but not without difficulty—and ultimately some students relied on alternative strategies to make a decision with respect to the rule, such as counting atoms which was similar to the mapping strategy identified previously.[36,44–47] With the app, however, students appeared to not consider pK_a or resonance. Many of these students began with simply trying mechanistic steps, using the app-directed tasks to guide their thinking. On the other hand, for the imidazole reaction, students in both groups knew to use the pK_a table to identify the specific site on the molecule where the reaction would occur, though they had difficulty utilizing the pK_a table as none of the exact structures from the reactions were present. This indicates that while students generally knew that they could use the pK_a table, they may not know how to effectively apply the information the pK_a table contains and may preferentially use it in lieu of chemical thinking. These findings align with the research by Flynn and Amellal who identified that students had difficulties using the pK_a table when given more complex molecules and when they needed to approximate pK_a values.[19]

Students from both groups frequently referred to resonance, aligning with findings by Ferguson and Bodner.[44] They demonstrated a range of thinking with respect to the resonance concept, many exhibiting learning difficulties similar to those described by Taber and Kim et al.[48,49] In the dicarbonyl reaction, students exhibited an understanding of the concepts, or rules, relating to resonance stabilization when determining the site where the reaction would occur. However, students' approach to the imidazole reaction revealed some difficulties with the syntax of resonance, where a number of students focused on resonance stabilization in the reactant rather than the product when determining the relative acidity of the two nitrogen atoms. This is similar to work by Cartrette and Mayo which indicates that students can identify the importance of resonance for assessing acidity or basicity, but may struggle to apply it successfully.[18] Furthermore, this ability to determine relative acidity is one of the ten necessary learning outcomes for the resonance concept as identified by Carle and Flynn.[50] Thus, it is valuable to recognize that not all students are meeting this learning outcome. A few students verbalized incorrect thinking about the relationships between resonance structures, specifically by expressing that various resonance structures are present as a mixture rather than contributing to the resonance hybrid. This incorrect understanding aligns with the previously reported findings that students consider resonance structures as distinct entities or as representations that denote rapid interconversion between double and single bonds.[48,49] As considering resonance structures can be important when determining how a reaction will proceed for many types of reactions,[50] it is key to build students' understanding of this concept and how to apply it in different contexts.

2.8.3 Students often considered one possible operation (i.e., mechanistic pathway), unless otherwise prompted

Our analysis indicates that there may be a difference between app and paper–pencil students in the extent to which they consider multiple mechanistic pathways. The paper–pencil students did not as often consider different possibilities in order to select the most likely mechanistic pathway and, for these students, incorrect decisions were often carried throughout the remainder of the reaction without notice or led to frustration later in the mechanism when they identified that something was not correct. This frustration compelled students to simply stop working on the reaction. On the other hand, students using the app were able to try different electron movements to see what the app would allow. The app students were able to get feedback from the app and could use this to guide their decision-making. This is not without drawback, as students tended to try things before considering the chemical feasibility of different possible mechanistic steps. However, some students did apply chemical reasoning after determining the mechanistic steps to explain why a particular step was correct once the app accepted the electron movements they tried. The app also prevented students from making and justifying incorrect mechanistic steps, providing targeted hints that could guide their thinking and constraining students from making chemically incorrect moves. This is particularly valuable in that it prevents students from the frustration caused by carrying through chemically infeasible steps that might lead students to stop thinking about the reaction altogether.

2.9 Limitations

There are a few limitations to this study inherent to the methodology used. This study was small and qualitative in nature and so the claims are limited in that we may not have captured the full range of students' thinking regarding acid–base reaction mechanisms and cannot make claims as to the relative prevalence of conceptions discussed herein. This study also only included students from a single institution and thus the results may not broadly apply across institutions. A larger sample size across a range of institutions may have revealed a greater range of conceptions and indicated differences in conceptions due to students' prior chemistry knowledge, the order in which the material is taught, and instructor methods. Specifically, most of the students at the study institution bypass general chemistry at the undergraduate level and go directly into first semester organic chemistry. Additionally, we might expect different reasoning by students who went through a revised curriculum such as that described by Flynn and Ogilvie.[51] While a quantitative study using survey methodology could provide information about the relative prevalence of students' conceptions, our study design was able to capture individualized conceptions. Additionally, while utilizing the two modalities allowed us to elicit a range of thinking across the students, there were inherent differences in the think-aloud procedures for the two groups of students that may have led to differences in student responses. However, in developing the interview protocol, and during the expert validation of the chosen reaction mechanisms, we attempted to ensure that the problem representation and prompting most aligned with how students would authentically engage with the different modalities, while mitigating differences from features other than the modalities and their inherent differences in prompting (e.g., providing both groups of students with pK_a tables).

2.10 Conclusions and implications

This study captured how students thought through acid–base reaction mechanisms by using two different modalities—i.e., paper–pencil and app based—and applied a models and modelling framework to examine the chemical features and concepts that students used to inform the mechanistic steps they made. Students' thinking was elicited through think-aloud interviews in which students worked through two acid–base reaction mechanisms either on paper or using the "Mechanisms" app. In general, students from both groups focused on the explicit features present in the modality they were using with minimal consideration of implicit electronics. They were familiar with the pertinent steps and rules for acid–base reactions, such as needing to determine the acidic and basic sites in a given reaction, and were familiar with the syntax used to make judgments about such rules, such as considering pK_a values or resonance. However, they often exhibited difficulty in applying the syntax to make decisions about the rules for the given reactions, indicating a poor conceptual grounding. Additionally, students showed reliance on explicit features, supports, and prompting—the nature of which differed between modalities—and did not always exhibit chemical thinking. For example, students resorted to strategies such as counting atoms to determine the acidity or basicity of a molecule, identifying similar structures on a pK_a table without thinking about implicit structural features, or using the app for guidance before using their own content knowledge. While resources such as the pK_a table or prompts provided by the app can be useful and support learning, it is important to train students to use these resources to support their critical thinking.

The results of this study have implications for both research and practice. Utilizing both the app and paper–pencil modalities for the think-aloud interviews elicited a greater range of student thinking. Therefore, this interview methodology has potential for future research focused on student thinking about reaction mechanisms and supports using multiple modalities to probe different thinking strategies that students may utilize. Our findings indicate that future research expanding this work to different reaction types or institutions may be merited. In particular, it would be valuable to compare students' thinking across institutions that use different instructional methods to teach the organic chemistry curriculum, such as that described by Flynn and Ogilvie.[51] Additionally, with the increased prevalence of app-based instructional tools, it is important to understand how these tools do or do not impact student thinking. Our results indicate that the app can be helpful for guiding student thinking and providing beneficial feedback to prevent students from performing chemically infeasible steps or obtaining incorrect products. However, additional scaffolding by instructors to promote reflective thinking may be necessary to mitigate rote use of the app. Promoting this type of reflective thinking would also benefit students working through reaction mechanisms in the traditional mode on paper, by helping them consider multiple reaction pathways and the chemical feasibility of proposed mechanistic steps.

2.11 Conflicts of interest

This research was funded in part by Alchemie, the company that created the "Mechanisms" app. Alchemie did not play any role in research design, data collection, or data analysis.

2.12 Appendices

2.12.1 Appendix 1. App goal cards and initial reaction screens

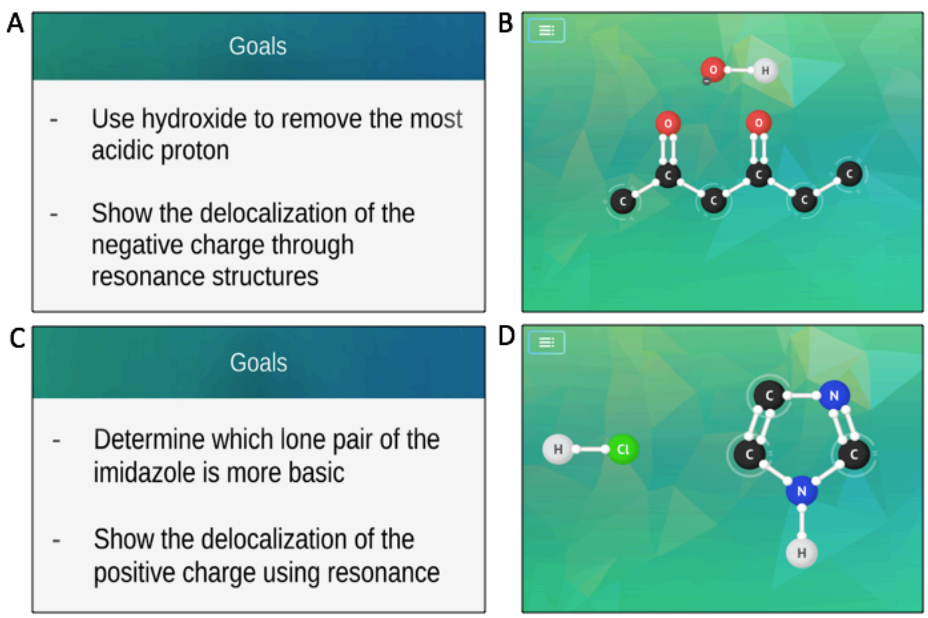


Figure 2.3 Goal cards (A and C) and initial reaction screens (B and D) seen by the app students as they worked through the 1,3-dicarbonyl and imidazole reactions, respectively.

2.12.2 Appendix 2. Excerpts from pK_a table

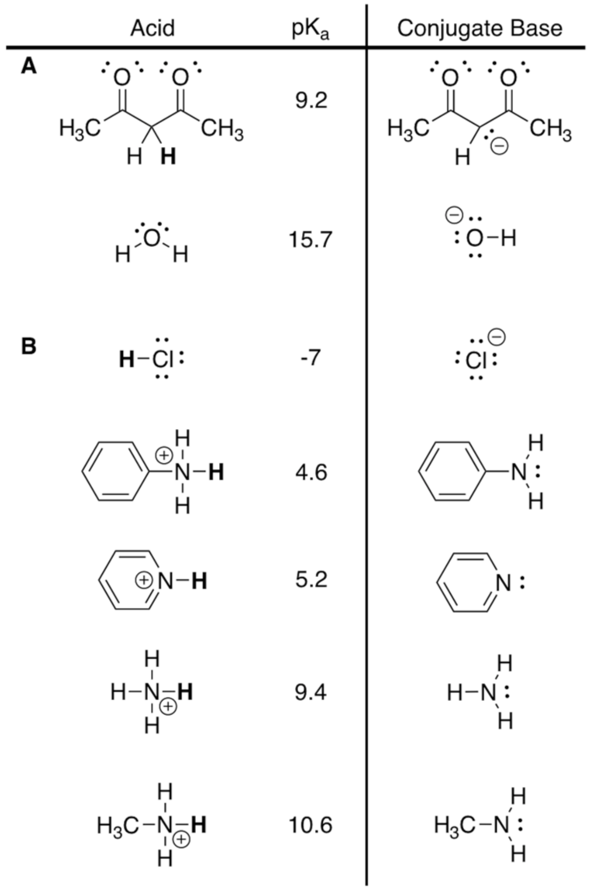


Figure 2.4 The structures that students referenced from the pK_a table they received during the think-aloud interviews: (A) pK_a values relevant to the 1,3-dicarbonyl reaction and (B) pK_a values relevant to the imidazole reaction. Students were provided with the complete pK_a table they use in the organic chemistry courses at the study institution.

2.12.3 Appendix 3. Coding scheme

Table 2.3 Coding scheme

Parent code	Sub-code	Definition	Exemplars
Chemical considerations	Protonation/deprotonation	Student discusses where protonation or deprotonation will occur or talks about protonating/deprotonating during a step of the reaction.	"This one's been protonated, it's going to take hydrogen from somewhere…"
	Acid–base	Student identifies the acid, base, or the acidic/basic site on a molecule or in the reaction.	"That's a strong acid that will dissociate. HCl…"
	Charge	Student thinks about the role charged atoms play in directing the reaction steps or discusses charge on atom/molecule. Charge can be implicitly mentioned (i.e., talking about further reaction at carbocation because it is unstable).	"I'm looking at this and I don't think carbon wants to have that negative charge very much."

	Carbocation	Student explicitly mentions a carbocation. This could be the presence of, formation of, or stabilization of a carbocation.	"Yes. Actually no. Because you can't really move the double bond around too much because then the carbon will become a carbocation."
	Resonance	Student talks about the presence of resonance structures or resonance stabilization.	"I know, in this, the resonances look different to me."
	Electronegativity	Student considers the electronegativity of various atoms to help determine reactivity.	"The oxygen's more electronegative, so that's going to be more likely to have that negative charge."
Reaction step	Bond breaking/forming	Student explicitly talks about breaking or forming a bond during the reaction step.	"I'll drag one of the electron pair to the hydrogen and break the hydrogen bond to form the water, and now we have a negatively charged carbon atom"
	Electrons	Student explicitly talks about electrons or lone pairs that are present or moving during the reaction step.	"…so this is allowed to move the electrons."
	Molecule/atom-focused	Student talks about a molecule or atom reacting during the reaction step.	"Alright. I know HCl is a really good acid, which means that it likes to give its hydrogen away."
Justification	Recognizes reaction component or step	Student recognizes a step/component of a reaction because they know it is a step/component of the type/classification of reaction they are doing. Often they explicitly identify some surface features to identify the step or type of reaction; this can be species in the reaction, functional groups, individual atoms, bonds, etc. (not just stating reaction type because this is told to them).	"so that tells me that this is a proton addition, or proton transfer, reaction."
	App hint/goal/task card directed action	Student explicitly verbalizes that a hint, goal, or task card directed their action.	"and then the arrows also showed the electrons that are this double bond over here to get the oxygen lone pairs."
Student actions	Incorrect	Student makes a move that is incorrect. Co-coded with the chemical feature/move that is incorrect.	"So, I'll drive one of the hydrogens to the oxygen. Not gonna work."

	Draw or pop out implicit protons or lone pairs	Student draws out the protons or lone pairs on a line-angle notation molecule; also code if they redraw molecules as Lewis structures.	"…okay. I'm gonna say it keeps this lone pair. Just ... all right. And then you have 1, 2, 3 C's and five Hs."
	Counting atoms	Student counts atoms at the beginning to identify what changes or at the end to make sure all atoms are accounted for.	"So this one, isopropyl formula, this one is two, three, four, five, six, C6 with two O's"
	pK_a table	Student references the pK_a table provided or verbalizes memorized pK_a values.	"To see if, well I know this is a strong acid but I see it's pK_a and see if it can protonate one of the two nitrogens"
App-specific	Hint	Student gets a hint during the puzzle.	"not the most basic lone pair... positive charges... resonance structures. Right, so. Yeah. I'm going to just restart."
	Goals	Student looks at the goals during the puzzle.	"it told me that wasn't the…"
	Trying random things	Student starts trying random actions to find something that will work.	"I don't even know what I'm trying to do at this point."
	Restarted puzzle	Student restarts the puzzle mid-reaction.	"And so, restart that."

2.13 Acknowledgements

This work has been supported, in part, through a grant to Alchemie from the National Science Foundation Small Business Innovation Research program, #1659983, and the Michigan Corporate Relations Network (MCRN), funded by the Michigan Economic Development Corporation (MEDC) and administered by the University of Michigan Business Engagement Center and the U-M Economic Growth Institute's Small Company Innovation Program (SCIP), #AWD006745. This work was also supported by the National Science Foundation Graduate Research Fellowship Program, #DGE1256260. Additionally, we would like to thank the students who participated in our study.

2.14 References

(1) Stoyanovich, C.; Gandhi, A.; Flynn, A. B. Acid-Base Learning Outcomes for Students in an Introductory Organic Chemistry Course. *J. Chem. Educ.* **2015**, *92* (2), 220–229. https://doi.org/10.1021/ed5003338.

(2) 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.

(3) Grove, N. P.; Cooper, M. M.; Cox, E. L. Does Mechanistic Thinking Improve Student Success in Organic Chemistry? *J. Chem. Educ.* **2012**, *89* (7), 850–853. https://doi.org/10.1021/ed200394d.

(4) Bell, E.; Provost, J.; Bell, J. K. Skills and Foundational Concepts for Biochemistry Students. *ACS Symp. Ser.* **2019**, *1337*, 65–109. https://doi.org/10.1021/bk-2019-1337.ch004.

(5) Cowie, J. M. G.; Arrighi, V. *Polymers: Chemistry and Physics of Modern Materials*, 3rd ed.; CRC Press: Boca Raton, FL, 2007.

(6) Grove, N. P.; Cooper, M. M.; Rush, K. M. Decorating with Arrows: Toward the Development of Representational Competence in Organic Chemistry. *J. Chem. Educ.* **2012**, *89* (7), 844–849. https://doi.org/10.1021/ed2003934.

(7) Cooper, M. M.; Grove, N. P.; Pargas, R.; Bryfczynski, S. P.; Gatlin, T. OrganicPad: An Interactive Freehand Drawing Application for Drawing Lewis Structures and the Development of Skills in Organic Chemistry. *Chem. Educ. Res. Pract.* **2009**, *10* (4), 296–301. https://doi.org/10.1039/b920835f.

(8) Cooper, M. M.; Grove, N.; Underwood, S. M.; Klymkowsky, M. W. Lost in Lewis Structures: An Investigation of Student Difficulties in Developing Representational Competence. *J. Chem. Educ.* **2010**, *87* (8), 869–874. https://doi.org/10.1021/ed900004y.

(9) Larson, B. Molecules. Sunset Lake Software 2012.

(10) Libman, D.; Huang, L. Chemistry on the Go: Review of Chemistry Apps on Smartphones. *J. Chem. Educ.* **2013**, *90* (3), 320–325. https://doi.org/10.1021/ed300329e.

(11) McCollum, B. M.; Regier, L.; Leong, J.; Simpson, S.; Sterner, S. The Effects of Using Touch-Screen Devices on Students' Molecular Visualization and Representational Competence Skills. *J. Chem. Educ.* **2014**, *91* (11), 1810–1817. https://doi.org/10.1021/ed400674v.

(12) Mechanisms. Alchemie Solutions, Inc. 2018.

(13) Duffy, P. L.; Enneking, K. M.; Gampp, T. W.; Amir Hakim, K.; Coleman, A. F.; Laforest, K. V.; Paulson, D. M.; Paulson, E. T.; Shepard, J. D.; Tiettmeyer, J. M.; et al. Form versus Function: A Comparison of Lewis Structure Drawing Tools and the Extraneous Cognitive Load They Induce. *J. Chem. Educ.* **2019**, *96* (2), 238–247. https://doi.org/10.1021/acs.jchemed.8b00574.

(14) McClary, L.; Talanquer, V. College Chemistry Students' Mental Models of Acids and Acid Strength. *J. Res. Sci. Teach.* **2011**, *48* (4), 396–413. https://doi.org/10.1002/tea.20407.

(15) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid-Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(16) Garnett, P. J.; Garnett, P. J.; Hackling, M. W. Students' Alternative Conceptions in Chemistry: A Review of Research and Implications for Teaching and Learning. *Stud. Sci. Educ.* **1995**, *25* (1), 69–96. https://doi.org/10.1080/03057269508560050.

(17) Anderson, T. L.; Bodner, G. M. What Can We Do about "Parker"? A Case Study of a Good Student Who Didn't "get" Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–101. https://doi.org/10.1039/b806223b.

(18) Cartrette, D. P.; Mayo, P. M. Students' Understanding of Acids/Bases in Organic Chemistry Contexts. *Chem. Educ. Res. Pract.* **2011**, *12* (1), 29–39. https://doi.org/10.1039/c1rp90005f.

(19) Flynn, A. B.; Amellal, D. G. Chemical Information Literacy: PKa Values-Where Do Students Go Wrong? *J. Chem. Educ.* **2016**, *93* (1), 39–45. https://doi.org/10.1021/acs.jchemed.5b00420.

(20) Bhattacharyya, G. Practitioner Development in Organic Chemistry: How Graduate Students Conceptualize Organic Acids. *Chem. Educ. Res. Pract.* **2006**, *7* (4), 240–247. https://doi.org/10.1039/B5RP90024G.

(21) Shaffer, A. A. Let Us Give Lewis Acid-Base Theory the Priority It Deserves. *J. Chem. Educ.* **2006**, *83* (12), 1746–1750. https://doi.org/10.1021/ed083p1746.

(22) Dood, A. J.; Fields, K. B.; Raker, J. R. Using Lexical Analysis to Predict Lewis Acid-Base Model Use in Responses to an Acid-Base Proton-Transfer Reaction. *J. Chem. Educ.* **2018**, *95* (8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177.

(23) Bhattacharyya, G. From Source to Sink: Mechanistic Reasoning Using the Electron-Pushing Formalism. *J. Chem. Educ.* **2013**, *90* (10), 1282–1289. https://doi.org/10.1021/ed300765k.

(24) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(25) Winter, J. E.; Wegwerth, S. E.; DeKorver, B. K.; Morsch, L. A.; DeSutter, D.; Goldman, L. M.; Reutenauer, L. M. The Mechanisms App and Platform: A New Game-Based Product for Learning Curved Arrow Notation. In *Active Learning in Organic Chemistry: Implementation and Analysis*; Houseknecht, J. B., Leontyev, A., Maloney, V. M., Welder, C. O., Eds.; American Chemical Society, 2019; pp 99–115. https://doi.org/10.1021/bk-2019-1336.ch007.

(26) Bongers, A.; Beauvoir, B.; Streja, N.; Northoff, G.; Flynn, A. B. Building Mental Models of a Reaction Mechanism: The Influence of Static and Animated Representations, Prior Knowledge, and Spatial Ability. *Chem. Educ. Res. Pract.* **2020**, *21* (2), 496–512. https://doi.org/10.1039/c9rp00198k.

(27) Lesh, R. A.; Hoover, M.; Hole, B.; Kelly, A.; Post, T. Principles for Developing Thought-Revealing Activities for Students and Teachers. In *Handbook of Research Design in Mathematics and Science Education*; Kelly, A., Lesh, R. A., Eds.; Routledge, 2000; pp 591–645.

(28) Briggs, M.; Bodner, G. A Model of Molecular Visualization. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer: Netherlands, 2005; pp 61–72.

(29) Briggs, M. W. Models and Modeling: A Theory of Learning. In *Theoretical Frameworks for Research in Chemistry/Science Education*; Bodner, G. M., Orgill, M., Eds.; Pearson Prentice Hall, 2007; pp 69–82.

(30) Cohen, L.; Manion, L.; Morrison, K. Sampling. In *Research Methods in Education*; Routledge, 2011; pp 143–164.

(31) Herrington, D. G.; Daubenmire, P. L. Using Interviews in CER Projects: Options, Considerations, and Limitations. *ACS Symp. Ser.* **2014**, *1166*, 31–59. https://doi.org/10.1021/bk-2014-1166.ch003.

(32) Ericsson, K. A.; Simon, H. A. Verbal Reports as Data. *Psychol. Rev.* **1980**, *87* (3), 215–251.

(33) Corbin, J. M.; Strauss, A. Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qual. Sociol.* **1990**, *13* (1), 3–21. https://doi.org/10.1007/BF00988593.

(34) Kirilenko, A. P.; Stepchenkova, S. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLoS One* **2016**, *11* (3), 1–14. https://doi.org/10.1371/journal.pone.0149787.

(35) McHugh, M. L. Lessons in Biostatistics Interrater Reliability : The Kappa Statistic. *Biochem. Medica* **2012**, *22* (3), 276–282.

(36) Galloway, K. R.; Stoyanovich, C.; Flynn, A. B. Students' Interpretations of Mechanistic Language in Organic Chemistry before Learning Reactions. *Chem. Educ. Res. Pract.* **2017**, *18* (2), 353–374. https://doi.org/10.1039/c6rp00231e.

(37) Bhattacharyya, G.; Harris, M. S. Compromised Structures: Verbal Descriptions of Mechanism Diagrams. *J. Chem. Educ.* **2018**, *95* (3), 366–375. https://doi.org/10.1021/acs.jchemed.7b00157.

(38) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402–1407. https://doi.org/10.1021/ed082p1402.

(39) Domin, D. S.; Al-Masum, M.; Mensah, J. Students' Categorizations of Organic Compounds. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 114–121. https://doi.org/10.1039/b806226a.

(40) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Ideas about Nucleophiles and Electrophiles: The Role of Charges and Mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797–810. https://doi.org/10.1039/c5rp00113g.

(41) Graulich, N.; Bhattacharyya, G. Investigating Students' Similarity Judgments in Organic Chemistry. *Chem. Educ. Res. Pract.* **2017**, *18* (4), 774–784. https://doi.org/10.1039/c7rp00055c.

(42) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/c8rp00131f.

(43) Graulich, N.; Hedtrich, S.; Harzenetter, R. Explicit: Versus Implicit Similarity-Exploring Relational Conceptual Understanding in Organic Chemistry. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 924–936. https://doi.org/10.1039/c9rp00054b.

(44) Ferguson, R.; Bodner, G. M. Making Sense of the Arrow-Pushing Formalism among Chemistry Majors Enrolled in Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 102–113. https://doi.org/10.1039/b806225k.

(45) Bhattacharyya, G. Trials and Tribulations: Student Approaches and Difficulties with Proposing Mechanisms Using the Electron-Pushing Formalism. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 594–609. https://doi.org/10.1039/c3rp00127j.

(46) 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.

(47) Webber, D. M.; Flynn, A. B. How Are Students Solving Familiar and Unfamiliar Organic Chemistry Mechanism Questions in a New Curriculum? *J. Chem. Educ.* **2018**, *95* (9), 1451–1467. https://doi.org/10.1021/acs.jchemed.8b00158.

(48) Taber, K. S. Compounding Quanta: Probing the Frontiers of Student Understanding of Molecular Orbitals. *Chem. Educ. Res. Pr.* **2002**, *3* (2), 159–173. https://doi.org/10.1039/b2rp90013k.

(49) Kim, T.; 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.

(50) 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.

(51) 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.

Chapter 3
Eliciting Students' Reasoning About Acyl Transfer Reactions With Case Comparisons

3.1 Initial remarks

This chapter presents the second of two studies that investigate different instructional supports to elicit organic chemistry students' reasoning about organic reaction mechanisms. While Chapter 2 demonstrated the use of a mobile device application to elicit students' reasoning, this chapter explores the use of a case comparison problem structure to elicit students' reasoning. This problem structure involves asking students to make a claim in response to a question that requires comparing between two similar reactions; the study specifically focused on using case comparison problems to elicit students' reasoning about acyl transfer reactions at three time points throughout the semester. This research explores how these problems can promote students' reasoning at each time point, and additionally examines how students' reasoning changes over time.

The study specifically used a case study approach to investigate three second-semester organic chemistry students' reasoning with the case comparison problems across the semester. Think-aloud interviews were conducted at two time points, and students' written responses were collected from a case comparison activity that occurred in-class between the two time points. Students' responses to both the think-aloud interviews and the in-class activity were qualitatively analyzed to identify the concepts and features of the reaction mechanism students were using in their reasoning. The analysis was guided by the resources framework, which serves as a model for interpreting how students use resources (i.e., ideas about a phenomenon) that are activated in a given situation to solve problems and construct explanations. The findings describe the variety of resources that students activated when solving the case comparison problem and detail how students used these resources in their problem solving. Many of the concepts and ideas that students activated were underlying chemical properties which aligned with the explicit differences between the electrophiles, nucleophiles, and products in the two reactions making up the case comparisons. When making a claim about the reactions, some students would incorrectly apply concepts or ideas while still arriving at the correct solution. Another key finding was that the students exhibited varying degrees of ability to consider multiple resources in their reasoning. For example, in the first set of interviews, one student reasoned using only once resource whereas another student reasoned by weighing multiple resources. While students exhibited differences in ability to weigh resources in their initial interviews, all students demonstrated the ability to activate and weigh multiple resources in the final set of interviews. Key implications extending from this study are that case comparison problems can be effective for supporting students' reasoning with underlying properties, but that it is important for instructors to elicit students' reasoning for how they arrive at a final solution. This can be an important practice that allows instructors to identify whether students' understanding of the fundamental concepts and ideas used in their reasoning aligns with the learning goals for the course.

This chapter was originally published as a research article in *Chemistry Education Research and Practice*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author, I contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing). I. Zaimi contributed to conceptualization, methodology, data collection, analysis, and writing (review and editing). D. Kranz contributed to conceptualization and writing (review and editing). N. Graulich and G.V. Shultz contributed to project supervision, conceptualization, and writing (review and editing).

<u>*Original publication and copyright information:*</u>

Reproduced from F.M. Watts, I. Zaimi, D. Kranz, N. Graulich, and G.V. Shultz, *Chem. Educ. Res. Pract.* 2021, **22**, 364–381 with permission from the Royal Society of Chemistry.

3.2 Abstract

Reasoning about organic chemistry reaction mechanisms requires engagement with multiple concepts and necessitates balancing the relative influence of different chemical properties. A goal of organic chemistry instruction is to support students with engaging in this type of reasoning. In this study, we describe our use of case comparison problems to elicit students' reasoning about acyl transfer reaction mechanisms across a semester. Using an instrumental case study methodology, we analyzed three students' reasoning across three time points: in a pre-

interview at the beginning of the semester, on their written responses to one implementation of an in-class scaffold activity, and in a post-interview near the middle of the semester. Through the theoretical lens of Hammer's resources framework, we analyzed the resources that students activated when approaching the case comparison problems. We characterized how students used each resource to support their reasoning, alongside characterizing how students weighed the different resources they activated. Our findings indicate that the case comparison problems activated a number of resources for each student across the time points by encouraging students to relate the surface-feature differences between reactions with the associated underlying properties. Students generally used resources, such as resonance and steric effects, in similar ways to support their reasoning across the time points. The study also illustrates the range in students' abilities to weigh multiple conceptual influences and how this ability might change across the semester. This case study has implications for future research exploring how students reason with multiple concepts and for instructors seeking to implement activities that support students' reasoning with case comparison problems.

3.3 Introduction

Learning organic chemistry requires students to engage with core conceptual ideas that connect a large number of different reaction types and mechanisms. Hence, students need to engage with learning strategies that promote process-oriented reasoning and problem-solving skills over product-oriented, rote learning.[1,2] However, research shows that students often approach learning in organic chemistry by systematically memorizing specific conceptual relationships, reactions, and mechanisms rather than using process-oriented understandings of conceptual ideas.[1,3–6] Therefore, it is necessary for researchers and instructors to understand and promote students' reasoning about mechanisms in organic chemistry. In this work, we describe a case study to explore students' reasoning for case comparisons, which are problems that involve posing a question alongside two similar mechanisms that have purposefully designed contrasting features. Herein, we present our analysis of students' abilities to consider and weigh different concepts for case comparison problems about acyl transfer reactions at three time points in a second-semester organic chemistry laboratory course.

3.3.1 Reasoning in organic chemistry

Practicing organic chemists use mechanisms as explanatory and predictive tools for describing how and why reactions occur.[7] Instruction regarding reaction mechanisms typically involves presenting the electron-pushing formalism for this purpose. However, research shows that many students do not necessarily use mechanisms as intended or understand the physical meaning associated with the electron-pushing formalism.[8–11] Furthermore, research demonstrates that students have challenges interpreting the underlying properties that are communicated by other representations, such as molecular structures, with a tendency to focus on their surface features.[4,5,12–17]

Alongside the evidence suggesting students' limited understanding of the electron-pushing formalism and other representations in organic chemistry, several studies provide evidence of students' conceptual understanding related to organic reaction mechanisms.[2] Recent studies focus on specific reaction types typically taught within organic chemistry, including acid–base reactions,[18–21] addition reactions,[22] substitution reactions,[14,23–27] and elimination reactions,[14,24] among other reaction types.[15,28–30] These researchers describe how students apply conceptual understanding to different tasks related to organic reaction mechanisms. In particular, these studies demonstrate the range in how students can apply their conceptual understanding of key concepts in organic chemistry (e.g., acid–base chemistry, resonance, nucleophilicity, etc.) differently across reaction types. Research of students' understanding of specific concepts, including charge, resonance, and nucleophilicity, specifically demonstrates that students tend to focus on the structural features of molecules over their function.[5,9,14,21,31] While many studies explore the range of concepts and understandings students consider, it is necessary to explicitly understand how students weigh multiple conceptual considerations in their reasoning about organic reaction mechanisms.

The existing research demonstrates that students have a range of understanding of core concepts in organic chemistry and that students often treat mechanism tasks as product-oriented exercises. Hence, there is an ongoing effort in the chemistry education research community to suggest better ways to teach organic reaction mechanisms. In particular, it is important to know how students use multiple concepts when considering problems involving mechanisms. Therefore, it is necessary to research approaches that elicit students' reasoning and support students as they engage with connecting underlying conceptual principles to mechanistic steps occurring during a reaction. It is particularly valuable to research approaches beyond traditional mechanisms tasks, such as predicting products or drawing mechanistic arrows, since such problems are not as effective at eliciting students' reasoning.[32] In this study, we use case comparison problems designed to support and encourage mechanistic reasoning and explore the development of students' abilities to consider and weigh multiple concepts during their reasoning.

3.3.2 Case comparisons to elicit students' reasoning

Case comparison problems better elicit students' reasoning as compared to problems involving a single case.[24,33] In recent work, Graulich and Schween describe the relationships between case comparisons and the epistemic practices of organic chemists and discuss how case comparison tasks can support students' abilities to develop, apply, and expand upon their conceptual understanding.[34] Caspari et al. demonstrate the usefulness of case comparisons for eliciting students' reasoning in an interview setting, with scaffolding questions that support students' construction of more complex explanations.[27] In Bodé et al.'s research, students demonstrate their reasoning on a case comparison exam question.[23] Notably, Bodé et al. found that students with more sophisticated reasoning made direct comparisons between structures.[23] However, many students did not make comparisons for all explicit and implicit features in the reactions.

Similarly, in a study using eye-tracking, Rodemer et al. found that students tended to focus their attention on the reactants instead of the products when solving case comparisons.[35] Rodemer et al. also found that advanced students were faster and had increased focus on relevant chemical structures compared to beginner students.[35] Together, these studies demonstrate the usefulness of case comparisons for eliciting students' reasoning and point to a need for further research into how students consider each part of a case comparison problem. It is necessary to understand how students use concepts when solving case comparisons, if the concepts students consider change during a semester, and how to elicit such reasoning in a classroom setting.

Prior research on case comparison problems for organic reaction mechanisms using an instructional scaffold is reported by Caspari and Graulich, who described interviews in which students compared activation energies for the leaving group departure step of similar E1 reactions.[24] The interviewer first asked students to reason about the problem without the scaffold, followed by asking students to complete the same problem using an instructional scaffold. The researchers designed this scaffold to help students engage in reasoning with multiple variables by separating structural differences from mechanistic changes and by delineating the different influences each structural difference has on each change. By asking students to complete the same case comparison problem with and without the scaffold, the researchers found that the scaffold successfully built upon the reasoning structures students exhibited without the scaffold. Furthermore, they found that students' use of the scaffold was correlated with an increase in the number of influences students considered in their reasoning. To build upon this work, our goal was to identify if students' consideration of multiple properties changes during a semester for similar case comparison tasks, including students' reasoning as presented on an in-class activity similar in structure to the scaffold used in the work by Caspari and Graulich.[24]

3.4 Theoretical framework

3.4.1 Hammer's resources framework

This research is guided by Hammer's resources framework that describes an approach towards understanding the cognitive structures people use to construct explanations.[36,37] This framework is influenced by previous literature seeking to define units of cognition, including diSessa's "phenomenological primitives" and "coordination classes" and Thagard's "propositions."[38–40] Within Hammer's framework, these fine-grained cognitive elements of knowledge are referred to as "resources."[36,37] Resources are, generally, ideas held about a phenomenon that are neither right nor wrong, and which are activated within certain situations to construct explanations. Activated resources can then be deemed productive or unproductive, depending on how the person relates the resources to the problem at hand. The resources framework contrasts with frameworks that suggest conceptions of a phenomenon are stable, fully formed ideas that are either correct or incorrect. Within these frameworks, conceptions are the cognitive units a person uses to build an explanation. Hammer suggests that conception-based frameworks do not accommodate the flexibility often observed in people's reasoning. That is, a conceptions framework does not explain situations in which a person seems to significantly alter a conception or misconception when encountering similar problems. Furthermore, the resources framework suggests that the resources activated when constructing explanations for similar phenomena differ depending on the situation, positing that the resources people activate for a particular phenomenon are not immutable. Since people use activated resources to construct explanations, the framework is useful for understanding how people's explanations for similar phenomena might change across time.

The resources framework provides a way for understanding how students construct explanations while engaging in organic chemistry case comparison reaction mechanism problems. Students have resources, or units of knowledge relating to structural features of a representation or concepts such as resonance, induction, or electronegativity, that they should be able to use to construct explanations. Students activate these resources for case comparison problems to explain the mechanistic question within the problem. The resources framework offers a way to understand what specific resources students activate and how they use ideas from activating multiple resources to produce explanations. Furthermore, the framework can be useful for understanding how students' reasoning changes, in terms of what resources a task activates at multiple time points and how students can use these resources when constructing explanations.

3.5 Research questions

This research aims to understand how students engage in reasoning and constructing explanations for case comparison problems about acyl transfer mechanism problems in organic chemistry. The goal is to qualitatively understand if and how students' reasoning for these problems might develop during the semester. This research addresses this goal by focusing on two aspects of students' reasoning for organic mechanism case comparison problems, namely what resources they activate and how they weigh activated resources across three time points. To address this goal, we seek to answer the following questions:

1. What resources do students activate when considering the case comparison problems and how do the resources students activate change across time?
2. How do students weigh resources when constructing explanations for the case comparison problems across time?

3.6 Methods

3.6.1 Instrumental case study methodology and research design

This research presents an exploratory, instrumental case study to investigate how students engage with case comparison problems during a second-semester introductory organic chemistry course for majors and non-majors at the University of Michigan. Instrumental case studies are those in which specific cases are used to understand a phenomenon, in contrast to traditional case studies that seek to understand something about the cases themselves.[41] In this research, we are studying the phenomenon of how students' engagement with case comparison problems develops throughout a second-semester organic chemistry laboratory course. Hence, this study aims not to understand the students themselves but to understand how their reasoning with case comparison problems develops. The research design involved collecting data at three times during the semester: a pre-interview at the beginning of the semester, completed worksheets from an in-class activity, and a post-interview in the weeks following the activity. Because students' reasoning is complex, particularly concerning how students activate and weigh multiple resources, the instrumental case study methodology is appropriate to provide a detailed, qualitative characterization of differences in how students respond to case comparison tasks. The case study methodology allows for a detailed analysis of how students reason with these problems across time points. Furthermore, the case study methodology is useful for guiding future research of students' reasoning on case comparison tasks.[41,42]

3.6.2 Setting

This study was conducted as part of a larger study at a research university in the Midwestern United States. The research was situated within the second-semester introductory organic chemistry laboratory course, which is offered separately from the lecture course at the study institution. The laboratory course consists of a weekly one-hour lecture taught by the course instructors and a four-hour laboratory taught by graduate student instructors. Students worked with case comparisons in different aspects of the course, including in the laboratory itself and on assignments. The first-semester lecture and laboratory courses are both prerequisites for the second-semester laboratory course, and the second-semester lecture course is an advisory prerequisite or co-requisite. Students usually take the introductory organic chemistry sequence in their first or second year, followed by inorganic, analytical, and physical chemistry courses in later years for chemistry majors.

3.6.3 Participants

This study is part of a larger research effort for which we recruited nine students to participate. We selected three participants, given the pseudonyms Brooke, Violet, and Chad, to focus on for the instrumental case study. These students were selected due to the comparative level of detail in their responses across time points. Furthermore, the observed similarities and differences between these participants during the data collection process presented a rich set of data for which it was valuable to employ the instrumental case study methodology. These specific students were also chosen because their reasoning was representative of the reasoning observed during the data collection process across all nine participants. By focusing on three students, we can provide a detailed account of how they responded to the case comparison problems across time points. The pseudonyms assigned to participants do not reflect their race, ethnicity, gender, or other identities.

3.6.4 Data collection

The data collected for this study includes pre-interviews, students' responses to an in-class activity, and post-interviews, across which students responded to two different case comparison problems. All data collection procedures received Institutional Review Board approval for human subjects research, and all students consented to participate in data collection. The semester was fourteen weeks long, and the pre-interviews took place in the second and third weeks, the in-class activity took place in the sixth week, and the post-interviews took place in the ninth and tenth weeks.

Pre- and post-interviews. For the data used in this study, all three pre-interviews and one post-interview were conducted in-person and audio recorded. Documents annotated by both the interviewer and interviewee were collected. Two post-interviews were conducted and recorded *via* video conferencing software, and documents were shared and annotated by participants using Google Drive. We conducted all interviews as think-aloud interviews, in which the interviewer instructed participants to verbalize their thinking as they considered the case comparison and responded to the guiding question described below.[43] Students were provided with a copy of a periodic table and pK_a table to use if needed during each interview.

Both interview protocols involved asking students to reason through a case comparison problem involving acyl transfer reactions, shown as presented to students in Figure 3.1 (A) and (B). We asked students to decide which of the mechanistic steps shown in the case comparison has the lowest activation energy. The acyl transfer reactions chosen for this case comparison were selected because they required considering two variables: the different substituent on the electrophile (methoxy *versus* chlorine) and the different nucleophiles (hydroxide *versus* methylamine). Through the interviews, we aimed to capture how students considered the two variables in the contrasting cases as they decided about the relative activation energies for the represented steps.

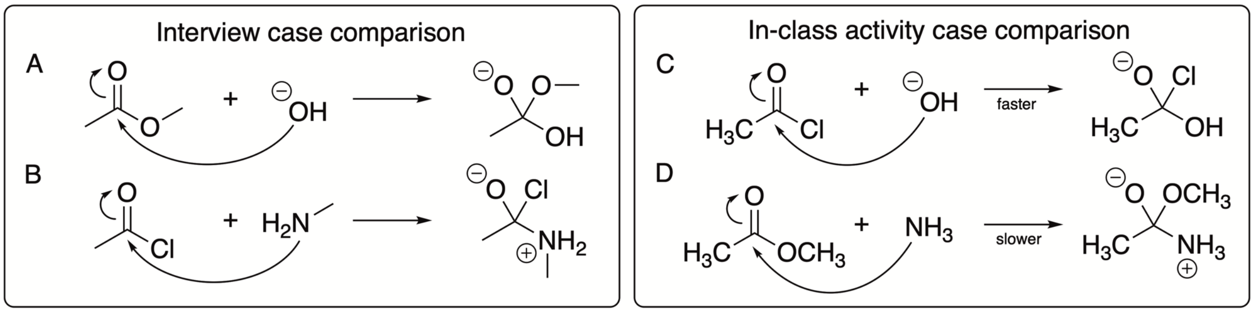


Figure 3.1 The case comparison reactions for the pre- and post-interviews (reactions A and B) and for the in-class activity (reactions C and D), shown as presented to the students.

The pre-interview protocol included a warm-up for the case comparison problem, in which the interviewer asked students to think-aloud while describing the two reactions. The interviewer asked probing questions to encourage students to identify similarities and differences between the reactions and describe the depicted mechanistic steps. After the warm-up, the interviewer asked students the guiding question, "Which reaction has a lower activation energy for the represented step? Make a prediction." Students were instructed to annotate the document as needed and to think-aloud as they responded to the question. Probing questions were asked after students completed their responses to clarify their intended meaning as well as their working definitions of activation energy and other concepts brought up during the task. Students appeared to have an appropriate working definition of the concept of activation energy. After the probing questions, the interviewer asked students to formulate a final statement of their response to the guiding question along with a summary of their reasoning. At the time of the pre-interview, students had not yet learned about acyl transfer reactions in their organic chemistry course but could reasonably be expected to respond to the problem based on the material covered earlier in the course sequence.

The post-interview protocol presented students with the same case comparison problem and asked the same guiding question from the pre-interview. Probing questions were asked to clarify students' statements, and then interviewers asked students to formulate a final statement and summary of their reasoning. At the time of the post-interview, students had experience with acyl transfer reactions both in the lecture and laboratory.

In-class activity and implementation. The in-class activity was developed based upon the activity described in prior research reported by Caspari and Graulich.[24] The development of the activity followed the process described in detail by Graulich and Caspari.[44] The activity took the form of a scaffold implemented during the one-hour lecture component of the course to guide students through reasoning about organic chemistry case comparison problems. The activity asked students to explain relative reaction rates for contrasting cases of single mechanistic steps. During the implementation, students worked in small groups to complete the activity as the instructor and multiple graduate student instructors circulated through the lecture hall to address student questions. Afterward, activity worksheets were collected from students, including the three participants' completed activities. The acyl transfer reactions in the activity, shown as presented to students in Figure 3.1 (C) and (D), were similar to those used in the pre-interviews in that the case comparison required students to consider two variables in their reasoning: the different substituents on the electrophile (chlorine *versus* methoxy) and the different nucleophiles (hydroxide *versus* ammonia). We collected this data as a mid-point in the data collection process to provide insight into how students' reasoning developed from pre- to post-interview. The data served as an artefact of students' reasoning for a problem similar to the case comparison problem in the pre- and post-interviews. The complete activity is reproduced in 3.11.1 Appendix 1.

3.6.5 Data analysis

Throughout the analysis, the research team focused on presenting what students were doing when responding to the case comparison problems rather than evaluating or assessing students' responses. As such, the three data sources were used to develop detailed, complete profiles of students' reasoning for each participant.[45] The interview recordings were transcribed verbatim, and we used the transcripts and recordings to write detailed descriptions of the resources students activated as they reasoned about the case comparison problems. Similarly, the in-class activity worksheets were scanned, and we used students' writings and annotations to write detailed descriptions of how students' reasoning was presented on the in-class activity. All profile descriptions incorporated annotations for how students used resources to guide their reasoning and how they weighed different resources when constructing explanations. These descriptions were read by another member of the research team and cross-referenced with the original data sources to ensure they accurately represented the students' reasoning.

The detailed descriptions of students' reasoning in each data source were then inductively coded by two research team members.[45] They independently analyzed the data for (1) the resources students activated when considering the problems and (2) how students weighed multiple resources, which we defined as the process of identifying resources to be more or less important for constructing an explanation in response to the guiding question. Afterward, the researchers discussed the inductive coding, organized the codes into themes, and developed finalized coding schemes.[45] The coding schemes are presented in 3.11.2 Appendix 2. With the final coding schemes, the two researchers then re-coded the data, discussed the coding, and reached a consensus on the final set of codes applied to each data source. The detailed descriptions of students' reasoning and the coding results were then used to develop more concise profiles to represent the students' reasoning, presented in 3.11.3 Appendix 3. We then used cross-case analysis, in which members of the research team discussed the coding results and profiles to make comparisons between students and across time points, to identify the key findings from the data. By using the cross-case analysis methodology, we were able to examine the similarities and differences across profiles to more deeply understand how the students approached the case comparison problems.[45] Discussions with the research team took place throughout the process of profile development, coding, and cross-case analysis to ensure the reliability and trustworthiness of the results.

3.7 Results and discussion

This study aims to understand how students approach case comparison problems focused on acyl transfer reactions during the second-semester organic laboratory course. The presented analysis focuses on the resources students activated when considering case comparison problems and how students weighed resources when constructing explanations. The profiles and coding that describe students' explanations on the case comparison problems in the pre-interviews, on an in-class activity, and in the post-interviews serve as the basis of the cross-case analysis for the results and discussion and are presented in the appendices. The presented analysis seeks to identify the similarities and differences for each student across time points and between students at each time point.

3.7.1 What resources do students activate when considering the case comparison problems and how do the resources students activate change across time?

To address this research question, we will first discuss the resources students considered across time points, which are directly tied to the differences that students observed between the electrophiles, nucleophiles, and products in the presented reactions. Then we will discuss how students' activated resources changed across the time points of the study. This analysis is drawn from the inductive coding for resources students used to guide their reasoning (with the detailed coding scheme presented in 3.11.2 Appendix 2) and the individual themes, as they emerged from the data, are presented across Table 3.1, Table 3.2, and Table 3.3.

Resources activated when considering differences between the electrophiles. The two electrophiles for both the interview and in-class activity case comparison problems were carboxylic acid derivatives, with only one difference between them: the chlorine of the acid chloride *versus* the methoxy group substituent of the ester, as shown in Figure 3.2. When considering this difference, all three students activated the resources of resonance, induction, and sterics at different points during the data collection. Only one student, Brooke, activated the resource of reactivity trends. When students activated each of these resources is summarized in Table 3.1.

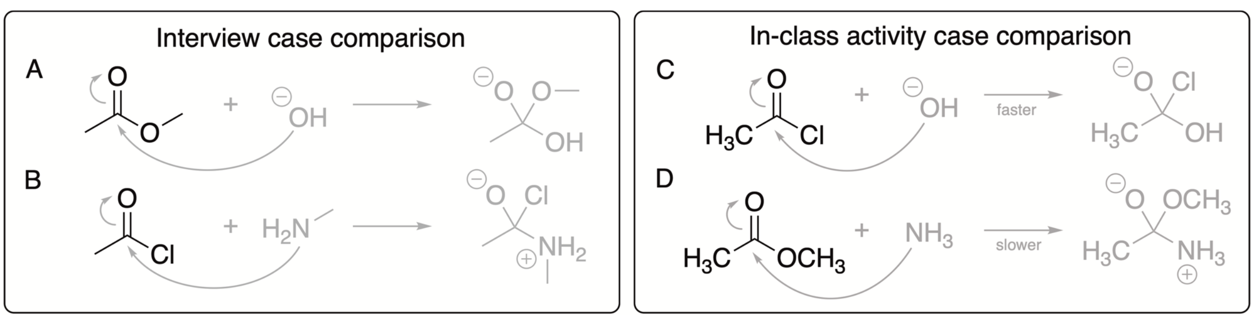


Figure 3.2 The case-comparison reactions for the pre- and post-interviews (reactions A and B) and for the in-class activity (reactions C and D), with the electrophiles emphasized.

Table 3.1 Resources students activated when considering the electrophile in the case-comparison problems. For each student, the checked (✓) boxes indicate the resources students activated at each time point (pre- interview, in-class activity, and post-interview). The crossed (X) boxes indicate resources the students did not activate

Resource	**Brooke**			**Violet**			**Chad**		
	Pre	In-class	Post	Pre	In-class	Post	Pre	In-class	Post
Resonance structures	✓	✓	X	X	X	✓	✓	✓	✓
Inductive effects	X	✓	X	X	✓	✓	X	✓	✓
Steric bulk	✓	X	✓	X	✓	X	✓	X	X
Reactivity trends	X	✓	✓	X	X	X	X	X	X

Students activated the concept of resonance by recognizing the nonbonding electron pairs in the functional groups. For example, in Brooke's pre-interview they identified that both reactions have atoms with nonbonding electron pairs adjacent to the carbonyl that "have resonance with the oxygen in the carbonyl" which "should have some stabilizing effect." Similarly, in the post-interview, Chad identified that the methoxy group allows for resonance with the carbonyl, which "makes the negative charge more spread out and more present in the carbonyl" and "less electrophilic." That students appealed to resonance aligns with prior research demonstrating students' reliance on the concept.[9,21,22] Notably, when the students considered resonance within the electrophiles, they did not exhibit the challenges students often have with resonance, such as describing resonance structures as distinct entities rather than contributors to a resonance hybrid.[21,22,31,46,47]

Students similarly activated the concepts of induction or sterics when comparing the chlorine and methoxy functional groups. On the in-class activity, Violet considered sterics and induction in tandem, stating,

> "Reaction [C] is less sterically hindered as it doesn't have a bulky CH_3 [*sic*] group and also has an electron-withdrawing group (Cl) that makes the electrophilic carbon site more favorable for attack."

The other students similarly used these concepts in their reasoning to suggest that the electron-withdrawing chlorine would increase the reaction rate (or lower activation energy) while the steric bulk of the methoxy group would decrease the reaction rate (or raise activation energy). The students' consideration of both steric and inductive effects within the electrophiles exemplifies how they may reason using both explicit and implicit structural features without necessarily needing to infer implicit electronic properties from the explicit features, a possibility suggested in prior research.[27] However, students' use of steric considerations aligns with prior studies of students applying this concept when other, electronic properties are more appropriate.[23]

Brooke activated the resource of reactivity trends, which involved discussing specific knowledge of which functional groups are generally more or less reactive. For instance, on the in-class activity, Brooke wrote that the acid chloride is "highly reactive" while the ester is "moderately reactive." This type of reasoning similarly appeared in Brooke's post-interview, in which they began by identifying that Reaction B has "the most reactive possible carboxylic acid derivative," a statement which they used to claim that the "energetics of this reaction are very favorable." This type of reasoning reflects the acyl compound reactivity trends students learn in the lecture course at the study institution, in which it is emphasized that acyl halides are the most reactive carboxylic acid derivatives and amides are the least reactive. Brooke's recollection of this trend could be reflective of students' rote memorization of rules in organic chemistry that has been documented in the literature.[1,4,14,48,49]

Resources activated when considering differences between the nucleophiles. The nucleophiles in the two case comparison problems were hydroxide and an amine. The amine was methylamine for the interview problem, whereas the amine was ammonia for the in-class activity, as illustrated in Figure 3.3. When considering the different nucleophiles in the case comparisons, students reasoned with resources including charge, sterics, basicity, electronegativity, and reactivity trends. The resources that students activated at each time are summarized in Table 3.2.

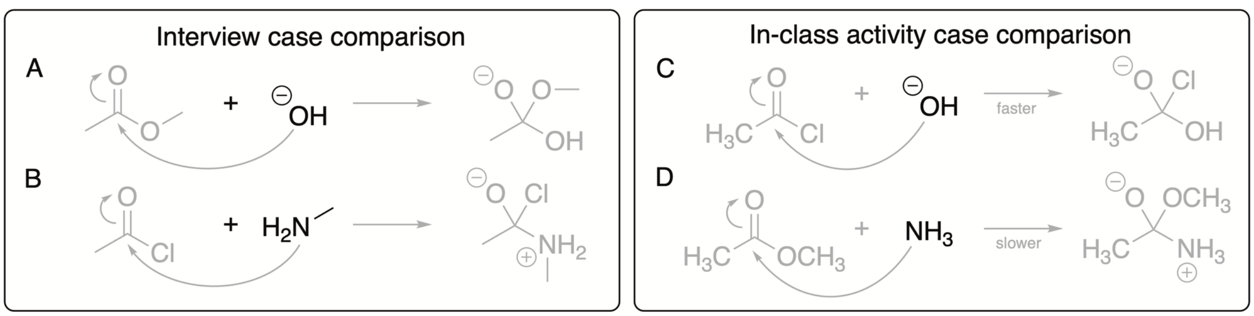


Figure 3.3 The case-comparison reactions for the pre- and post-interviews (reactions A and B) and for the in-class activity (reactions C and D), with the nucleophiles emphasized.

Table 3.2 Resources students activated when considering the nucleophile in the case-comparison problems. For each student, the checked (✓) boxes indicate the resources students activated at each time point (pre-interview, in-class activity, and post-interview). The crossed (X) boxes indicate resources the students did not activate

Resource	**Brooke**			**Violet**			**Chad**		
	Pre	In-class	Post	Pre	In-class	Post	Pre	In-class	Post
Formal charge	X	X	✓	✓	✓	✓	X	✓	✓
Steric bulk	✓	X	✓	X	X	X	✓	X	X
Basicity	X	X	X	X	✓	X	✓	X	X
Electronegativity	X	X	X	X	X	X	✓	X	X
Reactivity trends	X	X	✓	X	X	X	X	X	X

The most frequently activated resources were the charges of the nucleophile. When considering charges, students focused on the negatively charged hydroxide compared to the neutrally charged amine. For example, when considering the hydroxide in the pre-interview, Violet stated, “You’ve got a negative charge already present, which I know already would want to not have that. So it’s going to especially be attracted to getting rid of that negative charge.” Similarly, on the in-class activity, Chad wrote that the hydroxide “is a better nucleophile because [of the] negative charge.” In the post-interview, Brooke initially indicated similar reasoning, stating that their

> “immediate inclination is to think that the negatively charged species will be more reactive, but if I think about some trends here, I know that in general nitrogen species are going to be more nucleophilic than oxygen species.”

In Brooke’s case, they first noted a resource they did not use in their reasoning (the negative formal charge) which activated a resource they did use in their reasoning (the reactivity trends). Brooke was the only student who mentioned this reactivity trend for the nucleophiles, similarly to how Brooke was the only student to mention the reactivity trends for the electrophiles. In general, the students’ tendency to associate negative charges with nucleophilicity is similar to findings in the literature regarding how students conceptualize nucleophiles.[5,26,50]

Students also reasoned by comparing the relative steric bulk of the nucleophiles, such as in Chad’s pre-interview where they stated, “the nucleophile in Reaction [B] is more bulky and I think it’s going to have a harder time trying to attack,” suggesting that the increased steric bulk would correspond to higher activation energy for Reaction B. Brooke exhibited similar reasoning when considering sterics. Violet and Chad also activated resources relating to basicity when considering the nucleophile—exemplified by Violet writing “[hydroxide] strong base” and “[ammonia] weaker base” on their response to the in-class activity. Chad reasoned by using the provided pK_a table to identify that hydroxide “has the higher pK_a of its conjugate base,” connecting this to the fact that “[oxygen] is more electronegative” to identify that “[hydroxide] is going to be the better nucleophile” compared to methylamine. These students’ consideration of sterics and basicity is similar to previous work that has demonstrated students’ alignment of these concepts with nucleophilicity.[5,14]

Resources activated when considering differences between the product. The products for the case comparison reactions are emphasized in Figure 3.4. All students activated one particular resource, the differences between formal charges, across the data collection time points. One student, Violet, also activated the concept of resonance when comparing the products in the pre-interview. The resources students activated for the products and for which time point are summarized in Table 3.3.

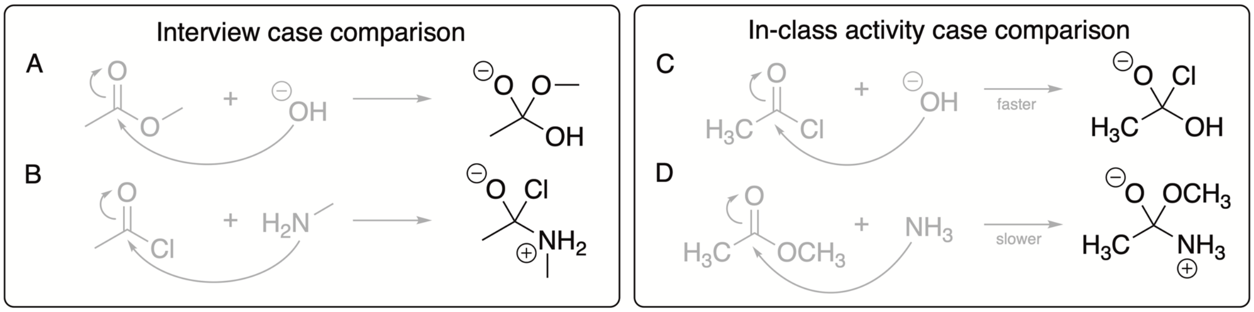


Figure 3.4 The case-comparison reactions for the pre- and post-interviews (reactions A and B) and for the in-class activity (reactions C and D), with the products emphasized.

Table 3.3 Resources students activated when considering the product in the case-comparison problems. For each student, the checked (✓) boxes indicate the resources students activated at each time point (pre-interview, in-class activity, and post-interview). The crossed (X) boxes indicate resources the student did not activate

Resource	**Brooke**			**Violet**			**Chad**		
	Pre	In-class	Post	Pre	In-class	Post	Pre	In-class	Post
Formal charge	✓	X	X	✓	✓	✓	✓	X	✓
Resonance structures	X	X	X	✓	X	X	X	X	X

All three students considered the charges in the product during the pre-interview. When using charges to reason, students attempted to make a connection between charges and stability. In Chad's pre- and post-interview, they reasoned about charges in the product by stating that

> "it seems like Reaction [B] is going to have a more stable product just because the overall net charge is zero, and in Reaction [A] the products has an overall net charge of minus one."

In contrast, Brooke and Violet inferred the opposite relationship between charge and stability. For example, Violet stated in the pre-interview that it "is not favorable, generally, to have two charges within a molecule" and wrote on the in-class activity that the product in Reaction [D] is "less stable (has more charges)." Similarly, Brooke stated in the pre-interview that "the charges in the final product [of Reaction A] are more favorable than the charges in the final product of Reaction [B]." All of the students used their reasoning about charges to make claims about which of the two products was more stable. They all reasoned that the more stable product would correspond to the reaction with the lower activation energy. Such connections between formal charges, stability, and activation energy align with prior research demonstrating how students reason about charges focusing on the products when asked about the activation energy of a mechanistic step.[27]

Only one student, Violet, considered resonance in the products, when neither product structure is capable of resonance stabilization. When comparing the reactions, Violet initially indicated that the product in Reaction A, but not the product of Reaction B, was capable of resonance stabilization: "There's also resonance stabilization in the molecule afterwards. I don't think there could be for Reaction [B]…" However, Violet goes on to realize the incorrect application of this resource, stating, "Well actually wait… No, I don't think you could do that… So maybe you wouldn't have resonance stabilization, which would put a slight problem in my theory." This was the only case of a student activating a resource and then recognizing that it would be conceptually incorrect to apply the resource. As identified by Carle and Flynn's research on learning objectives for resonance, the ability to recognize molecular structures able to engage in resonance stabilization is the first learning objective for the resonance concept, whereas using delocalization concepts to explain reactivity are among the final learning objectives.[51] That Violet considered resonance but did not identify the correct molecules in the reactions in which the concept applies aligns with students' application of the concept in prior studies.[21] Furthermore, Violet's reasoning suggests that their ability to use resonance as an explanatory concept is at an early stage in Carle and Flynn's proposed set of resonance learning outcomes—though it is important to note that Violet recognized that the concept, as they initially applied it, was not correct.[51]

Changes in activated resources across time points. For each student, the resources activated when examining the differences in the case comparisons changed from pre-interview to in-class activity to post-interview, as seen across Table 3.1, Table 3.2, and Table 3.3. The different resources students activated on the in-class activity may be tied to the slight differences in the framing of the problem—specifically, that the in-class activity problem identified the faster reaction and asked students to explain, whereas the interview problem asked the student to identify which reaction had the lower activation energy and explain their reasoning. Furthermore, there was an inherent difference in the setting between the in-class activity and the interviews. Hence, the comparisons made between the resources activated in the pre- and post-interviews are the most insightful, as the framing remained the same across these time points. However, due to the similarities between the reactions on the in-class activity and in the interviews, noting the resources students activated on the in-class activity is useful as a midpoint for gaining insight into how students' reasoning changed across the semester.

In the pre-interview, Brooke considered resonance and sterics for the electrophile; sterics for the nucleophile; and charge for the product. After activating these resources, Brooke selected Reaction A by focusing on charges. In the post-interview, Brooke activated sterics and reactivity trends for the electrophile and charge, sterics, and reactivity trends for the nucleophile. In contrast to the pre-interview, Brooke selected Reaction B by focusing on the reactivity trends in the electrophiles. A commonality in the resources Brooke activated from the pre- to post-interview was their discussion of sterics, which they described as having a possible influence on reactivity. However, in both interviews, they ultimately decided that sterics was not a factor that would change their response to the guiding question. Brooke was the only student to discuss reactivity trends, and most of the other resources Brooke considered for these time points (e.g., charges, resonance, and inductive effects) were related to how Brooke described these trends. The increased focus across time on the reactivity trends for the electrophile is evident from Brooke's response to the in-class activity, in which they only activated resources for the electrophile. Brooke's tendency in both the in-class activity and post-interview to activate resources focusing only on the electrophile or nucleophile aligns with prior research that identifies students' focus on reactants over products.[35]

In Violet's pre-interview, the only resources they activated were charges for the nucleophile and product, and—incorrectly—resonance for the product. That Violet did not activate any resources when considering the electrophile aligns with prior studies suggesting students do not necessarily focus on every feature when examining case comparison reactions.[23,35] Because Violet recognized that considering resonance in the product was incorrect, they selected Reaction A by only considering one resource: charges. For Violet's post-interview, they activated resonance structures and inductive effects in the electrophile and formal charges in the nucleophile and product. To respond to the guiding question, Violet used their consideration of resonance and induction for the electrophiles to select Reaction B. That Violet activated more resources from pre- to post-interview suggests a development in the ability to activate more resources in response to the problem. This possible development is also evident in their response to the in-class activity, where they activated more resources than they did for the pre-interview. Violet's reasoning about charges from pre- to post-interview was similar, but the fact that they considered more resources to ultimately select Reaction B in the post-interview demonstrates how the increase in activated resources shaped Violet's reasoning.

In Chad's pre-interview, they activated the resources of resonance and sterics for the electrophile; sterics, basicity, and electronegativity for the nucleophile; and charge for the product. They were unsure which resource to focus on to respond to the guiding question but selected Reaction A as their final response. In Chad's post-interview, they activated resonance and induction for the electrophile and charges for the nucleophile and product. They selected Reaction B by focusing on inductive effects in the electrophiles and the charges of the products. The resources Chad activated during the in-class activity were nearly the same as those activated during the post-interview, demonstrating how Chad narrowed the number of resources they activated from pre-interview to in-class activity to post-interview. This narrowing of activated resources was possibly valuable for Chad because they may have been unable to respond to the guiding question in the pre-interview due to the number of resources they activated. Two of the resources Chad activated in the post-interview were the same as those activated in the pre-interview—resonance in the electrophile and charges. While Chad used these resources similarly for the two time points, their more focused consideration of inductive effects in the electrophile guided their reasoning in the post-interview.

3.7.2 How do students weigh resources when constructing explanations for the case comparison problems across time?

We address this research question by focusing on how students weighed between multiple resources when constructing their explanations. As described above and in the coding scheme in 3.11.2 Appendix 2, we specify weighing resources as the process of students identifying which of the multiple resources they deem relevant to be more or less important in their reasoning. We examine how students engaged in weighing resources differently between students and across time points, as captured by the second layer of coding completed in the analysis process. The results of this analysis are presented in Table 3.4. In characterizing each student's response holistically for how they weighed resources, we seek to illustrate how students organized the resources activated by the case comparison problems and the in-class activity when constructing their responses.

Table 3.4 Coding scheme for characterizing how each student weighed resources when constructing their explanations. For each student, the checked (✓) boxes indicate the type of weighing students engaged in at each time point (pre-interview, in-class activity, and post-interview). The crossed (X) boxes indicate the type of weighing students did not engage in

Demonstrating weighing of resources	Brooke			Violet			Chad		
	Pre	In-class	Post	Pre	In-class	Post	Pre	In-class	Post
Uses one resource	X	X	X	✓	X	X	X	X	X
Does not weigh resources	X	X	X	X	✓	X	✓	✓	X
Weighs resources	✓	✓	✓	X	X	✓	X	X	✓

Students exhibited a range of abilities to weigh resources in the pre-interview. In the pre-interview, all participants selected Reaction A as having the lower activation energy. However, the way students weighed resources differed. Brooke weighed their considerations of resonance and charge, noting that resonance is "not the most important thing to consider" and focusing on the charges in the product to select Reaction A. Later in the interview, Brooke also exhibited placing less importance on their consideration of sterics, stating that the sterics are "going to have some impact, but I don't think it's going to be game-changing in this case." Violet activated two resources in their pre-interview: charges in the nucleophile and product, and, incorrectly, resonance in the product. Because Violet recognized that considering resonance stabilization within the product was incorrect, as described previously, Violet selected Reaction A based solely on their consideration of charges. Violet emphasized their focus on this resource, stating, "I always try to keep track of charges as best I can because that I find really helps me." Because Violet only considered one resource in making their decision, they did not weigh resources. Both Violet and Brooke's focus on charges in their explanations, despite their differences in considering and weighing other resources, aligns with prior research demonstrating students' reliance on formal charges when considering mechanisms.[5,15,27,29]

Chad activated the most resources of the three students during the pre-interview. While Chad activated many relevant resources, the number of resources they considered proved to be challenging for them. They used these considerations to alternately support Reaction A (when comparing nucleophiles) and Reaction B (when comparing electrophiles and products). Chad changed their mind frequently throughout the interview, explicitly recognizing that they "keep going back and forth." Near the end of the interview, Chad stated that "I think I'm just going to have to stick with the charges" to select Reaction B, before changing their mind a final time and selecting Reaction A due to the nucleophile and the resonance delocalization of the electrophile in

Reaction A. Ultimately, Chad exhibited considering resources which support different conclusions and explicitly indicated difficulty in selecting which resources to weigh as most important in constructing their response.

While all students activated at least two resources in the pre-interview, they demonstrated differences in how they weighed resources. Brooke weighed between multiple resources and ultimately focused their reasoning on charges in the product as the most important resource. Violet similarly focused on charges in the product to make a decision, only after recognizing their incorrect consideration of resonance in the products. In contrast to Violet, Chad activated many resources and demonstrated difficulty with weighing resources. While Violet was able to make a decision and Chad was not, it was likely that Violet's decision-making was possible because Violet only considered a single, relevant resource—a reasoning strategy that aligns with students' one-reason decision-making, as identified in the literature.[52] In contrast, Brooke and Chad considered multiple resources, with only Brooke exhibiting the ability to weigh between resources to make a decision. However, Chad's engagement in a productive struggle—in contrast to Violet's consideration of a single resource and apparent lack of difficulty in making a final decision—could have been useful for Chad's learning.

Students all considered multiple resources but did not necessarily demonstrate evidence of weighing resources on the in-class activity. On the in-class activity, students were prompted to explain why one of the shown reactions was faster than the other after considering all relevant properties. The three participants exhibited similarities in terms of the number of resources activated. However, there was a difference in whether students weighed resources. For the in-class activity, students demonstrated weighing resources by activating resources that would support different claims about which reaction was faster but ultimately building their explanation by focusing on the specific resources that support the claim for one of the reactions to be faster.

Brooke, who engaged in weighing resources during the pre-interview, also weighed resources on the in-class activity. This consistency in weighing was evident from how Brooke activated resources that supported different claims: the inductive effects of the chlorine and the resonance effects of the methoxy group, and the resource of general reactivity trends. Brooke indicated that the inductive effects of the chlorine would increase the rate of Reaction C but make the "double bond harder to break," whereas the resonance effects of the methoxy group would discourage Reaction D despite making the "double bond easier to break." By presenting reasoning with these resources that support either reaction being faster, Brooke demonstrated evidence of weighing the inductive effects of the chlorine over the resonance effects of the methoxy group.

While Brooke weighed resources on the in-class activity, Violet and Chad did not demonstrate evidence of weighing resources. They both only wrote about resources that supported the same conclusion and thereby did not necessarily weigh between resources that would have supported different conclusions. They each activated different resources, with some overlap. However, all of the resources that both Violet and Chad considered were used to support their explanation for why Reaction C was faster. As such, both students did not make visible any considerations that would provide a counterargument for Reaction D being faster, in the way that Brooke did. Hence, while Violet and Chad considered multiple resources, they did not produce evidence of weighing resources. However, it is possible that these students did activate resources that they weighed as less important by not including them in their written response. This could be an artefact of the activity itself—both the prompting on the activity and that it took place in-class rather than in an interview setting—as prior research suggests that prompt changes can influence students' exhibited mechanistic reasoning.[18,25]

Students exhibited convergence in ability to weigh resources in the post-interview. In the post-interview, all participants selected Reaction B as having the lower activation energy. Additionally, all students engaged in explicitly weighing resources. In Brooke's reasoning, they identified how each activated resource would support different reactions having lower activation energy. Similarly to their response during the pre-interview, Brooke explicitly identified that they "don't think that sterics are going to play a huge role." Brooke placed the most weight on reactivity trends of the electrophiles and nucleophiles in selecting Reaction B. However, given Brooke's abilities to weigh resources across the semester, it could be likely that Brooke's appeal to reactivity trends represents the use of the resource as an explanatory concept—i.e., a resource that is a collection of appropriate resources, such as resonance and inductive effects, that are responsible for determining reactivity trends.

Violet activated more resources on the post-interview as compared to their pre-interview. They explicitly identified that the electrophilicity of the acid chloride in Reaction B outweighs differences in charges on the nucleophiles: "Even though the nitrogen is not negatively charged like the oxygen, [the acid chloride] still makes [the nitrogen] better to possibly be attacking that [carbonyl carbon]." Violet also placed less emphasis on their consideration of charges in the products, selecting Reaction B despite stating that the products have "a negative charge and a positive charge in [the] product, which isn't great." Ultimately, Violet selected Reaction B after considering inductive and resonance effects in the electrophiles and charges in the nucleophiles and products, placing most emphasis on the differences between the electrophiles.

Chad activated a similar number of resources as the other participants during the post-interview. After considering each difference between the electrophiles to support Reaction B, Chad considered how the differences between the nucleophiles would support Reaction A. Then, when considering the products, Chad stated that "I still think actually Reaction [B] is going to be faster or have the lower activation energies, because I think the products also play a role." In concluding the interview, Chad reiterated their choice of Reaction B by weighing the inductive effects in the electrophile and the overall neutral charge of the product over the different nucleophile strengths.

Across the post-interviews, students activated a similar number of resources and demonstrated similar abilities to weigh which resources are the most important. Furthermore, Brooke and Violet emphasized their considerations of the electrophiles over the resources activated when considering the nucleophiles or products, with Brooke focusing on reactivity trends and Violet on inductive and resonance effects. Chad similarly weighed inductive effects in the electrophile alongside charges in the product as the most important resources. Ultimately, while the students activated and focused on slightly different resources in the post-interview, they all successfully engaged in weighing resources to construct their explanations.

3.8 Limitations

The primary limitation of this research stems from the case study methodology. While the instrumental case study allows for a thorough investigation and description of how the students engaged with the organic case comparison problems across the three data collection time points, it inherently limits the scope of claims that can be made from the study. While the three case study participants selected for analysis were representative of the range in reasoning observed across the nine participants during the data collection process, the findings reported in this study are not meant to be generalizable to larger populations of students or students at different institutional settings with different backgrounds and experiences. Additionally, the case study participants were recruited on a voluntary basis, possibly contributing to self-selection bias. Because of the methodological limitations, this research is also limited in the claims that can be made. While this study identified a broad range of resources activated across the case study participants, we cannot claim whether this demonstrated range captures all possible variability in students' reasoning. In particular, the resources activated on the in-class activity may have been influenced by the inherent differences between the activity and the interviews. Thus, the discussion of students' reasoning on the in-class activity is limited to identifying how their reasoning presented itself on the activity to provide context for the development in students' reasoning observed from the pre- to post-interviews.

3.9 Conclusions

This case study provides an analysis of three second-semester organic chemistry students' reasoning for acyl transfer case comparison problems across three time points: a pre-interview, an in-class activity, and a post-interview. The analysis of students' reasoning focused specifically on the resources students activated when considering these problems and how students weighed the different resources. Our results demonstrate how case comparison problems can elicit multiple resources, both in the interview setting and on the in-class activity. Furthermore, our findings indicate a range of students' abilities to engage in weighing resources.

Students activated a variety of resources for each stage of data collection, and the resources were not necessarily uniform across time points or between students. When students activated the same resources, they tended to reason similarly, both for each student and across time points. That is, when students considered the concepts of resonance, induction, and sterics, for example, they tended to use these concepts to support their explanations in similar ways. The findings regarding students' use of resources contributes to the literature on students' reasoning in organic chemistry by identifying the concepts students use when considering case comparisons of acyl transfer reactions. In particular, all resources students considered were directly related to the differences they identified between the case comparison reactions. That is, each resource tied directly to the explicit differences between the electrophiles, nucleophiles, and products in the reactions.

Students' abilities to weigh resources ranged from basing their explanations on one activated resource to explicitly considering multiple resources and making their decisions based on what they deemed to be the most important. One student also exhibited the ability to identify an incorrect resource that they activated. When considering multiple resources, students did not necessarily weigh these resources in the pre-interview or during the in-class activity. This tendency was present when students indicated not knowing how to balance the resources they were considering and when students only considered resources that supported the same conclusion. Students also demonstrated explicitly weighing resources by stating the importance of certain resources over others when providing their explanations. Most notably, students' activated resources and ability to weigh them differed between students and changed over time. While students began at different abilities, their ability to activate and weigh multiple resources converged over the semester.

3.10 Implications

3.10.1 Implications for research

Findings from this study indicate how students activate and weigh different resources when producing explanations for case comparison problems. However, future research is merited for furthering our understanding of how in-class activities can support students' reasoning. For instance, some of the differences observed in students' responses to the in-class activity and the interviews suggest the need for further research into how the framing of a prompt may influence the resources students activate and if they weigh resources. Furthermore, this research did not specifically examine the processes by which students deemed particular resources to be more productive than others, which is worth further study. The research design for this project—using pre- and post-interviews and artefacts of students' reasoning from an in-class activity— could also be extended to activities implemented with other instructors or at other institutions to increase the generalizability of the results presented herein. This exploratory case study focused on providing a detailed analysis of the resources students' activated and indicates a range in students' abilities to weigh resources across time points. Future research could seek to identify variations and nuances across students who demonstrate different reasoning abilities. In particular, the characterization of how students weighed resources in this study can be extended and applied to students' reasoning across the organic chemistry curriculum and into graduate programs. Future research could also develop and use quantitative measures of students' reasoning abilities to measure the effects of in-class activities on students' reasoning. Additionally, there is a need for further research exploring how instructors can use similar in-class activities to elicit and scaffold students' reasoning—i.e., both the activation and weighing of multiple resources—with the use of case comparison problems in classrooms on a larger scale. For example, it would be valuable to research modifications of the instructional scaffold used in this study to explore how the prompting might better support students' engagement in weighing multiple resources.

3.10.2 Implications for practice

This research includes an in-class scaffold activity and demonstrates how it can elicit students' reasoning for case comparison problems within a large lecture. In particular, the case comparison problem in this study was found to be particularly useful for focusing students' reasoning on all of the differences between reactants and supporting their identification of numerous explicit and implicit structural properties to guide their explanations. Furthermore, over the data collection period, our results illustrate that students were able to converge in their ability to weigh multiple resources, suggesting that students starting at different stages of reasoning ability can improve over a relatively short period of time.

Our findings indicate that, while students in general might consider a variety of concepts, specific students can consider different concepts for similar problems while arriving at the same answers. Furthermore, our results demonstrate that how students weigh concepts can differ. Together, these findings suggest incorporating instructional practices, such as the task design recently reported by Lieber and Graulich,[53] that both elicit and assess how students reason rather than focusing on the product of students' reasoning. For example, the students in this case study all provided the same final answer at each time point of data collection but differed in how they arrived at the answer. Hence, to support students' reasoning rather than ability to arrive at an answer, instructional practices must engage in eliciting, supporting, and assessing students' reasoning itself rather than the product of students' reasoning.

Other implications for practice relate to considerations specifically for teaching reaction mechanisms. We demonstrated how students might have incorrect or different understandings of fundamental topics in organic chemistry at the beginning of a second-semester, introductory organic chemistry course taken by both chemistry majors and non-majors. For example, one student within the case study indicated potential challenges with knowing what structures are capable of resonance stabilization; additionally, our participants exhibited different understandings of the relationship between charges in a molecule and relative stability. Hence, it is necessary for instructors in the middle of an organic chemistry course sequence to identify the core concepts from earlier semesters for which students may still need support in learning. Lastly, our results indicate a range in students' abilities to consider and weigh different resources across the semester. Instructors can use this finding to inform how they model reasoning strategies while connecting to students' existing problem-solving skills during instruction.

3.11 Appendices

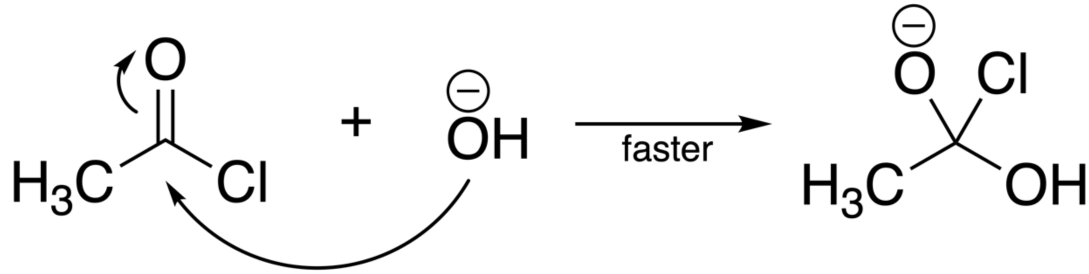
3.11.1 Appendix 1. The in-class activity

The three pages of the in-class activity worksheet are presented below in Figure 3.5, Figure 3.6, and Figure 3.7.

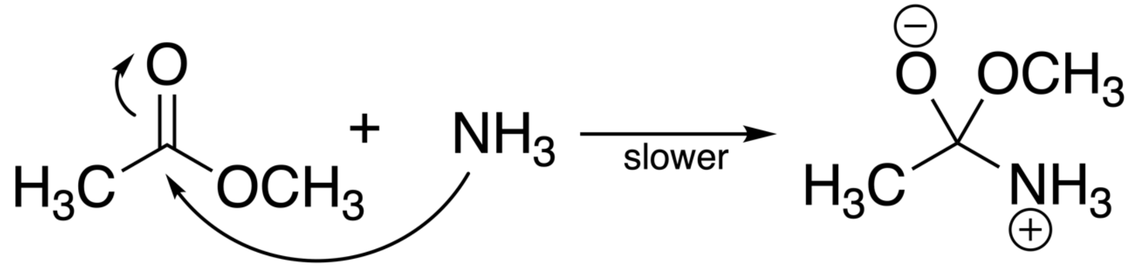
Name:

Why do reactions A and B occur at different speeds? Make a prediction. Think about explaining your prediction by making a claim and backing that claim up with evidence and reasoning.

A



B



Address the points below. Use the table on page two to organize your thoughts and the space on page three to write your final answer. The final answer should be an in-depth explanation at least a paragraph.

1. What structures differ in both reactions A and B? Specify the functional groups in which the reactants differ. Note these differences in the first row of the table (boxes 1).
2. What chemical and physical properties do the functional groups in (boxes 1) have? Note the properties in the second row of the table (boxes 2).
3. What changes occur from reactants to products in both reactions A and B? Focus on the functional groups specified in (boxes 1). Note the changes, such as forming a charge, breaking a bond, or making a bond, in the first column of the table (boxes 3).
4. Why do the changes in (boxes 3) occur? Consider concepts from an energy viewpoint or drawing a reaction coordinate diagram for both reactions A and B.
5. Describe as precisely as possible how the properties in (boxes 1) and (boxes 2) influence the property changes in (boxes 3). Do the influences of the properties accelerate the reaction step or slow it down? Do they have no effect at all compared to the other reaction? Note the influences in the middle boxes (boxes 4/5).
6. Provide a statement that answers the question: **why do reactions A and B occur at different speeds?**

Figure 3.5 The first page of the in-class activity.

Reminder: The numbers on the boxes correspond with the numbers from the instructions on the first page.

	DIFFERENCES		
	(**1**) atoms and functional groups A	vs.	(**1**) atoms and functional groups B
	(**2**) physical and chemical properties of atoms and functional groups A	vs.	(**2**) physical and chemical properties of atoms and functional groups B
SIMILARITES			
(**3**) change 1	(**4/5**) influence of properties A on change 1		(**4/5**) influence of properties B on change 1
(**3**) change 2	(**4/5**) influence of properties A on change 2		(**4/5**) influence of properties B on change 2

Figure 3.6 The second page of the in-class activity.

(6) **Why do reactions A and B occur at different speeds?** If stuck, consider using this sentence stem: "Reaction ______ occurs at a faster speed because ______ affects ______ by ______."

Figure 3.7 The third page of the in-class activity. The remainder of the page left space for students' responses.

3.11.2 Appendix 2. The coding schemes

The two coding schemes are presented in Table 3.5.

Table 3.5 The coding schemes for the analysis, including the resources students activated when considering the electrophiles, nucleophiles, and products, and the characterization of how each student weighed resources when constructing their explanations. Definitions and exemplars are provided for the three themes centered around resources students activated. Definitions are given for the characterization of how students weighed resources. As applying these codes required holistic evaluation of each data source, the profile descriptions corresponding to each code are indicated

Code	**Definition**	**Profiles/exemplars**
Coding scheme for resources activated when considering the electrophiles		

Resonance structures	The student considers the resonance structures or electron donating effects *via* resonance.	*Brooke, in-class activity:* "resonance effects make double bond easier to break"
Inductive effects	The student considers the inductive, electron-withdrawing effects of substituents.	*Violet, post-interview:* "Both of the groups on that first reagent are electron-withdrawing, which makes that carbon super partially positive."
Steric bulk	The student considers the relative sterics of substituents.	*Chad, pre-interview:* "The methoxy group is going to have more sterics than the chlorine group."
Reactivity trends	The student considers remembered reactivity trends for acyl compounds.	*Brooke, post-interview*: "I have the most reactive possible carboxylic acid derivative."
Coding scheme for resources activated when considering the nucleophiles		
Formal charge	The student considers the charge of the nucleophiles.	*Chad, in-class activity:* "[hydroxide] is a better nucleophile because negative charge"
Steric bulk	The student considers the relative sterics of the nucleophile.	*Brooke, pre-interview:* "There are some steric considerations here, as well. The hydroxyl group is not as bulky as the group in Reaction [B]."
Basicity	The student considers the basicity of the nucleophiles or the pK_a values of the nucleophiles' conjugate acids.	*Violet, in-class activity*: "[hydroxide] strong base" and "[ammonia] weaker base"
Electronegativity	The student considers the electronegativity of atoms in the nucleophiles.	*Chad, pre-interview*: "Also, [oxygen] is more electronegative so maybe that plays a role into it as well"
Reactivity trends	The student considers remembered reactivity trends for nucleophiles.	*Brooke, post-interview:* "If I think about some trends here, I know that in general nitrogen species are going to be more nucleophilic than oxygen species."
Coding scheme for resources activated when considering the products		
Charge	The student considers the charges of the products.	*Chad, post-interview:* "Because the products, you have an overall neutral charge for reaction [B] and for reaction [A] you have a negative charge."
Resonance structures	The student considers the possibility of resonance stabilization.	*Violet, pre-interview:* "There's also resonance stabilization in the molecule afterwards."
Coding scheme for characterizing how each student weighed resources when constructing their explanations		
Uses one resource	The student makes their decision based on one resource and thereby does not engage in weighing resources.	Violet, pre-interview
Does not weigh resources	The student considers two or more resources that support alternative conclusions, and explicitly indicates not knowing which resource(s) have more (or less) influence on their decision-making OR the student	Violet, in-class activity; Chad, pre-interview; Chad, in-class activity

	considers two or more resources that support the same conclusion, thereby suggesting that no resource has more (or less) influence on their decision-making.	
Weighs resources	The student considers two or more resources that support alternative conclusions, and suggests that at least one resource has more (or less) influence on their decision-making.	Brooke, pre-interview; Brooke, in-class activity; Brooke, post-interview; Violet, post-interview; Chad, post-interview

3.11.3 Appendix 3. Profile descriptions for each case

Case 1. Brooke

Pre-interview. Brooke begins by noting similarities in the bonds being broken and formed in the presented mechanism and noting differences in functional groups between the reactions. In response to the guiding question, Brooke considers the resonance effects in the electrophile then the charges on the products. Brooke weighs these two considerations, noting that resonance is "not the most important thing to consider" and focusing on charges to select Reaction A because it does not form charges in the product. After selecting Reaction A, Brooke mentions considering the sterics of the nucleophiles and electrophiles but states that the sterics are "going to have some impact, but I don't think it's going to be game-changing in this case." Ultimately, Brooke considers resonance in the electrophiles, the charges of the products, and sterics of all reactants but weighs charges most in selecting Reaction A.

In-class activity. Brooke identifies all the differences in the reactions and indicates properties relating to the electrophiles, writing about the electron withdrawing properties of the chlorine *versus* methoxy substituents and that the acid chloride is "highly reactive" while the ester is "moderately reactive." Brooke identifies that the electron withdrawing chlorine increases the reaction rate but makes the carbonyl pi-bond harder to break. For the ester, Brooke identifies that resonance effects discourage the reaction at the carbonyl carbon but that resonance makes the carbonyl pi-bond easier to break. Brooke ultimately explains that Reaction C is faster, though they presented reasoning that would support selecting either reaction. Hence, they implicitly weighed the electron withdrawing property of chlorine to be more important than the resonance effects of the methoxy group.

Post-interview. Brooke begins by identifying that Reaction B has "the most reactive possible carboxylic acid derivative" to claim that the "energetics of this reaction are very favorable." Brooke then considers the possible steric influence of methylamine before selecting Reaction B. Next, Brooke discusses the nucleophiles, stating that "the negatively charged species will be more reactive" but "in general nitrogen species are going to be more nucleophilic," and later changing their mind to state that "OH might be a little bit more reactive, but I don't think it's going to be by a huge amount." Brooke also considers the steric influences of the electrophiles but states that they "don't think that sterics are going to play a huge role." To conclude, Brooke reiterates selecting Reaction B after weighing the reactivity trends of the electrophiles more heavily than the reactivity of the nucleophiles or possible steric influences.

Case 2. Violet

Pre-interview. Violet begins by noting the differences in functional groups and focusing on the charges present in the reactions, stating "I always try to keep track of charges as best I can because that I find really helps me." In response to the guiding question, Violet immediately chooses Reaction A and then justifies their choice with their consideration of charges on the nucleophiles: "You've got a negative charge already present, which I know already would want to not have that. So it's going to especially be attracted to getting rid of that negative charge." They then consider the possibility for resonance stabilization within the product but realize their mistake in this reasoning because the products do not have delocalizable electron pairs, stating "so maybe you wouldn't have resonance stabilization, which would put a slight problem in my theory." After recognizing that it was incorrect to consider resonance, Violet concludes by reiterating their choice of Reaction A "because it's less charges in the first place." By only considering one resource, Violet does not engage in weighing between resources.

In-class activity. Violet identifies all differences between the reactions, indicating the steric and inductive effects influencing the electrophiles and the relative basicity of the nucleophiles. Violet indicates that the nucleophile in Reaction C will "attack faster" and that the product of Reaction C is "likely an intermediate and less reversible." They also indicate that the product in Reaction C will continue to react and form a more stabilized product, while the product in Reaction D is more likely to reverse to the starting materials due to the presence of more charges. Violet also states that Reaction C is faster "because of sterics and electron density," stating that the less sterically hindered and more electron withdrawing chlorine (*versus* the methyl group) increases the reaction rate. They also indicate the negative charge on the hydroxide to support Reaction C being faster. Each line of reasoning that Violet considers is in support of Reaction C. That is, Violet does not write about considerations that would support Reaction D and thus does not engage in weighing the importance of different resources.

Post-interview. Violet answers the guiding question by identifying differences in the reactions and considering the resonance effects possible in the electrophile for Reaction A that make the carbonyl carbon "less partially positive" and "less likely to be attacked." They also consider the inductive effects of the chlorine that increase the electrophilicity of the carbonyl carbon in Reaction B. They voice that the electrophilicity of the acid chloride in Reaction B outweighs differences in charges on the nucleophiles: "Even though the nitrogen is not negatively charged like the oxygen, [the acid chloride] still makes [the nitrogen] better to possibly be attacking that [carbonyl carbon]." Violet next considers the charges of the products, stating that the Reaction B products have "a negative charge and a positive charge in [the] product, which isn't great." To conclude, Violet selects Reaction B after considering inductive and resonance effects in the electrophiles and charges in the nucleophiles and products, placing most emphasis on the differences between the electrophiles.

Case 3. Chad

Pre-interview. Chad identifies the differences between the electrophiles and nucleophiles in the reactions and, in response to the guiding question, select Reaction A. They support their choice using the pK_a table and electronegativity to suggest that hydroxide is a better nucleophile. Chad then considers the resonance and sterics of the electrophile in Reaction A and the steric differences between the nucleophiles, considerations which they recognize support Reaction B. Then Chad states being conflicted about their response, keeping Reaction A as their answer, before identifying the charges in the products. Chad changes their mind to Reaction B because it "is going to have a more stable product just because the overall net charge is zero." They continue recognizing the conflicts in their thinking, specifically stating that "the nucleophile in Reaction [B] is more bulky and I think it's going to have a harder time trying to attack." Chad recognizes that they "keep going back and forth, but I think I'm just going to have to stick with the charges" and their selection of Reaction B, before again considering the resonance delocalization of the electrophile and better nucleophile in Reaction A. In the end, Chad selects Reaction A after considering a number of resources and struggling to weigh which resources are the most important.

In-class activity. Chad identifies all differences between the reactions and considers a number of properties: the inductive effects of the electronegative chlorine in Reaction C's electrophile, the resonance effects in Reaction D's electrophile, and the charges and relative strength of the nucleophiles. Chad writes that the charge on the hydroxide in Reaction C "makes it more nucleophilic and willing to react," while the ammonia is "a more neutral charge, that will be more stable and less likely to react as fast." Chad also indicates the influences of the differences between the electrophiles: that the carbon in Reaction C is "more electrophilic" while the carbonyl in Reaction D is resonance stabilized which makes the carbon "less electrophilic." Each influence Chad considers supports Reaction C, and thus Chad does not engage in weighing different resources.

Post-interview. Chad answers the guiding question by selecting Reaction B and supporting their claim by identifying that the electron withdrawing group on the electrophile makes the carbonyl more electrophilic. They identify the resonance effects in Reaction A's electrophile, which "makes the negative charge more spread out and more present in the carbonyl" and "less electrophilic." Chad then considers the different nucleophiles, stating that Reaction A has a stronger nucleophile because it has a negative charge, but despite this they suggest they are still in favor of Reaction B. Chad next considers the overall charges in the products, claiming that the overall neutral charge in the products of Reaction B further support their choice. To conclude, Chad reiterates their choice of Reaction B, weighing the inductive effects in the electrophile and the overall neutral charge of the product over the different nucleophile strengths.

3.12 Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260. The authors would additionally like to thank Solaire Finkenstaedt-Quinn, Nicholas Garza, Jessica Scott, Chang-Hwa Chiang, Maiya Yu, and members of the Shultz group for assistance with facilitating the in-class activity and discussions related to the preparation of this manuscript.

3.13 References

(1) Grove, N. P.; Bretz, S. L. A Continuum of Learning: From Rote Memorization to Meaningful Learning in Organic Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 201–208. https://doi.org/10.1039/c1rp90069b.

(2) Graulich, N. The Tip of the Iceberg in Organic Chemistry Classes: How Do Students Deal with the Invisible? *Chem. Educ. Res. Pract.* **2015**, *16* (1), 9–21. https://doi.org/10.1039/c4rp00165f.

(3) Anderson, T. L.; Bodner, G. M. What Can We Do about "Parker"? A Case Study of a Good Student Who Didn't "get" Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–

101. https://doi.org/10.1039/b806223b.

(4) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable Reasoning: Multi-Variate Problem-Solving in Organic Chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281–292. https://doi.org/10.1039/c0rp90003f.

(5) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Ideas about Nucleophiles and Electrophiles: The Role of Charges and Mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797–810. https://doi.org/10.1039/c5rp00113g.

(6) Webber, D. M.; Flynn, A. B. How Are Students Solving Familiar and Unfamiliar Organic Chemistry Mechanism Questions in a New Curriculum? *J. Chem. Educ.* **2018**, *95* (9), 1451–1467. https://doi.org/10.1021/acs.jchemed.8b00158.

(7) Goodwin, W. Mechanisms and Chemical Reaction. In *Philosophy of Chemistry*; Hendry, R. F., Needham, P., Woody, A. I., Eds.; Elsevier BV, 2012; Vol. 6, pp 309–327. https://doi.org/10.1016/B978-0-444-51675-6.50023-2.

(8) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402–1407. https://doi.org/10.1021/ed082p1402.

(9) Ferguson, R.; Bodner, G. M. Making Sense of the Arrow-Pushing Formalism among Chemistry Majors Enrolled in Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 102–113. https://doi.org/10.1039/b806225k.

(10) Grove, N. P.; Cooper, M. M.; Cox, E. L. Does Mechanistic Thinking Improve Student Success in Organic Chemistry? *J. Chem. Educ.* **2012**, *89* (7), 850–853. https://doi.org/10.1021/ed200394d.

(11) Grove, N. P.; Cooper, M. M.; Rush, K. M. Decorating with Arrows: Toward the Development of Representational Competence in Organic Chemistry. *J. Chem. Educ.* **2012**, *89* (7), 844–849. https://doi.org/10.1021/ed2003934.

(12) Domin, D. S.; Al-Masum, M.; Mensah, J. Students' Categorizations of Organic Compounds. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 114–121. https://doi.org/10.1039/b806226a.

(13) McClary, L.; Talanquer, V. College Chemistry Students' Mental Models of Acids and Acid Strength. *J. Res. Sci. Teach.* **2011**, *48* (4), 396–413. https://doi.org/10.1002/tea.20407.

(14) Cruz-Ramírez De Arellano, D.; Towns, M. H. Students' Understanding of Alkyl Halide Reactions in Undergraduate Organic Chemistry. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 501–515. https://doi.org/10.1039/c3rp00089c.

(15) Galloway, K. R.; Stoyanovich, C.; Flynn, A. B. Students' Interpretations of Mechanistic

Language in Organic Chemistry before Learning Reactions. *Chem. Educ. Res. Pract.* **2017**, *18* (2), 353–374. https://doi.org/10.1039/c6rp00231e.

(16) Galloway, K. R.; Leung, M. W.; Flynn, A. B. Patterns of Reactions: A Card Sort Task to Investigate Students' Organization of Organic Chemistry Reactions. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 30–52. https://doi.org/10.1039/c8rp00120k.

(17) Graulich, N.; Bhattacharyya, G. Investigating Students' Similarity Judgments in Organic Chemistry. *Chem. Educ. Res. Pract.* **2017**, *18* (4), 774–784. https://doi.org/10.1039/c7rp00055c.

(18) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid-Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(19) Dood, A. J.; Fields, K. B.; Raker, J. R. Using Lexical Analysis to Predict Lewis Acid-Base Model Use in Responses to an Acid-Base Proton-Transfer Reaction. *J. Chem. Educ.* **2018**, *95* (8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177.

(20) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(21) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.; Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions via App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 878–892. https://doi.org/10.1039/c9rp00260j.

(22) Finkenstaedt-Quinn, S. A.; Watts, F. M.; Petterson, M. N.; Archer, S. R.; Snyder-White, E. P.; Shultz, G. V. Exploring Student Thinking about Addition Reactions. *J. Chem. Educ.* **2020**, *97* (7), 1852–1862. https://doi.org/10.1021/acs.jchemed.0c00141.

(23) 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.

(24) Caspari, I.; Graulich, N. Scaffolding the Structure of Organic Chemistry Students' Multivariate Comparative Mechanistic Reasoning. *Int. J. Phys. Chem. Educ.* **2019**, *11* (2), 31–43. https://doi.org/10.12973/ijpce/211359.

(25) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(26) Dood, A. J.; Dood, J. C.; Cruz-Ramírez De Arellano, D.; Fields, K. B.; Raker, J. R. Analyzing Explanations of Substitution Reactions Using Lexical Analysis and Logistic Regression Techniques. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 267–286. https://doi.org/10.1039/c9rp00148d.

(27) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/c8rp00131f.

(28) Bhattacharyya, G.; Harris, M. S. Compromised Structures: Verbal Descriptions of Mechanism Diagrams. *J. Chem. Educ.* **2018**, *95* (3), 366–375. https://doi.org/10.1021/acs.jchemed.7b00157.

(29) Watts, F.; Schmidt-McCormack, J.; Wilhelm, C.; Karlin, A.; Sattar, A.; Thompson, B.; Gere, A. R.; Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21*, 1148–1172. https://doi.org/10.1039/C9RP00185A.

(30) Caspari, I.; Weinrich, M. L.; Sevian, H.; Graulich, N. This Mechanistic Step Is "Productive": Organic Chemistry Students' Backward-Oriented Reasoning. *Chem. Educ. Res. Pract.* **2018**, *19* (1), 42–59. https://doi.org/10.1039/c7rp00124j.

(31) 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.

(32) DeCocq, V.; Bhattacharyya, G. TMI (Too Much Information)! Effects of given Information on Organic Chemistry Students' Approaches to Solving Mechanism Tasks. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 213–228. https://doi.org/10.1039/c8rp00214b.

(33) Alfieri, L.; Nokes-Malach, T. J.; Schunn, C. D. Learning Through Case Comparisons: A Meta-Analytic Review. *Educ. Psychol.* **2013**, *48* (2), 87–113. https://doi.org/10.1080/00461520.2013.775712.

(34) Graulich, N.; Schween, M. Concept-Oriented Task Design: Making Purposeful Case Comparisons in Organic Chemistry. *J. Chem. Educ.* **2018**, *95* (3), 376–383. https://doi.org/10.1021/acs.jchemed.7b00672.

(35) Rodemer, M.; Eckhard, J.; Graulich, N.; Bernholt, S. Decoding Case Comparisons in Organic Chemistry: Eye-Tracking Students ' Visual Behavior. **2020**. https://doi.org/10.1021/acs.jchemed.0c00418.

(36) Hammer, D.; Elby, A. Tapping Epistemological Resources for Learning Physics. *J. Learn. Sci.* **2003**, *12* (1), 53–90. https://doi.org/10.1207/S15327809JLS1201.

(37) Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F. Resources, Framing, and Transfer. In *Transfer of Learning: Research and Perspectives*; Mestre, J., Ed.; Information Age Publishing: Greenwich, CT, 2004.

(38) diSessa, A. A. Toward an Epistemology of Physics. *Cogn. Instr.* **1993**, *10* (2–3), 105–225. https://doi.org/10.1080/07370008.1985.9649008.

(39) diSessa, A. A.; Sherin, B. L. What Changes in Conceptual Change? *Int. J. Sci. Educ.* **1998**, *20* (10), 1155–1191. https://doi.org/10.1080/0950069980201002.

(40) Thagard, P. Explanatory Coherence. *Behav. Brain Sci.* **1989**, *12*, 435–502. https://doi.org/10.4324/9780429494260-5.

(41) Grandy, G. Instrumental Case Study. In *Encyclopedia of Case Study Research*; Mills, A. J., Durepos, G., Wiebe, E., Eds.; SAGE Publications, Inc., 2012; pp 474–476. https://doi.org/:http://dx.doi.org.ezproxy.auckland.ac.nz/10.4135/9781412957397.

(42) Yin, R. *Case Study Research: Design and Methods*; SAGE Publications: London, 2014.

(43) Herrington, D. G.; Daubenmire, P. L. Using Interviews in CER Projects: Options, Considerations, and Limitations. *ACS Symp. Ser.* **2014**, *1166*, 31–59. https://doi.org/10.1021/bk-2014-1166.ch003.

(44) Graulich, N.; Caspari, I. Designing a Scaffold for Mechanistic Reasoning in Organic Chemistry. *Chem. Teach. Int.* **2020**, *0* (0). https://doi.org/10.1515/cti-2020-0001.

(45) Miles, M. B.; Huberman, A. M.; Saldana, J. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed.; Sage: Los Angeles, CA, 2014.

(46) Taber, K. S. Compounding Quanta: Probing the Frontiers of Student Understanding of Molecular Orbitals. *Chem. Educ. Res. Pr.* **2002**, *3* (2), 159–173. https://doi.org/10.1039/b2rp90013k.

(47) Kim, T.; 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.

(48) Christian, K.; Talanquer, V. Modes of Reasoning in Self-Initiated Study Groups in Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 286–295. https://doi.org/10.1039/c2rp20010d.

(49) Weinrich, M. L.; Talanquer, V. Mapping Students' Modes of Reasoning When Thinking about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 394–406. https://doi.org/10.1039/c5rp00208g.

(50) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Fragmented Ideas about the

Structure and Function of Nucleophiles and Electrophiles: A Concept Map Analysis. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 1019–1029. https://doi.org/10.1039/c6rp00111d.

(51) 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.

(52) Maeyer, J.; Talanquer, V. Making Predictions about Chemical Reactivity: Assumptions and Heuristics. *J. Res. Sci. Teach.* **2013**, *50* (6), 748–767. https://doi.org/10.1002/tea.21092.

(53) Lieber, L.; Graulich, N. Thinking in Alternatives—A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry. *J. Chem. Educ.* **2020**. https://doi.org/10.1021/acs.jchemed.0c00248.

Chapter 4
Writing-To-Learn To Support Engagement in STEM Courses

4.1 Initial remarks

This chapter focuses broadly on writing-to-learn (WTL) pedagogy as an instructional approach for supporting students' conceptual learning in STEM courses. Building on Chapters 2 and 3, which describe two instructional approaches for eliciting and supporting students' reasoning in organic chemistry, the remainder of the dissertation focuses on WTL pedagogy as another approach. As such, this chapter serves as a transition into the remainder of the dissertation. The chapter specifically focuses on the body of research extending from the MWrite program at the University of Michigan, which was established to work with STEM faculty to implement and evaluate evidence-based WTL pedagogy. The chapter reviews the sixteen research articles that explore students' learning experiences with WTL assignments implemented in MWrite courses. This research contributes to the literature on WTL by describing the types of learning supported by the evidence-based WTL assignment design promoted through the MWrite program. Furthermore, the study demonstrates that evidence-based design and implementation practices can allow for effective incorporation of writing to support students' learning in STEM courses.

The results of the sixteen research articles are synthesized and analyzed through the lens of a four-dimensional framework for engagement, which defines engagement as a phenomenon that spans cognitive, behavioral, emotional, and social domains. The first two key findings from the analysis use evidence from multiple MWrite studies to indicate the utility of WTL assignments for supporting students' abilities to describe content and for engaging students in disciplinary thinking practices. For example, many of the reviewed studies include evidence (through analyzing students' writing or revisions) that students can successfully describe the content targeted by WTL assignments. Similarly, results from some of the studies indicate the utility of WTL assignments to engage students in disciplinary thinking practices, such as engaging in argumentation from evidence or reasoning about mechanisms. The findings also include evidence from various studies pertaining to how students' affective experiences are influenced by the structure of MWrite assignments. Specifically, studies across course contexts demonstrate how the rhetorical components of WTL assignments—such as the genres, contexts, and audiences in the assignment descriptions—promote positive affective learning experiences by connecting to students' interests and building their motivation for learning the content. The final key finding focuses on results across studies demonstrating how the peer review and revision components of MWrite WTL assignments provide sources of knowledge for students that lead to content-focused revisions. For example, many of the studies indicate that students can provide content-focused peer review comments, which, combined with the drafts students read when providing feedback, influence the nature of students' revisions. By synthesizing the influences of the evidence-based design principles behind MWrite on students' engagement, the results provide several implications about WTL assignment design and implementation for instructors seeking to implement WTL in their courses.

This chapter will be published as a research article in a forthcoming issue of the *International Journal for the Scholarship of Teaching & Learning*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. Both myself and S.A. Finkenstaedt-Quinn are primary authors on the manuscript and contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing). A.R. Gere and G.V. Shultz contributed to funding acquisition, project supervision, conceptualization, and writing (review and editing).

4.2 Abstract

The writing-to-learn (WTL) literature is varied in how assignments are structured and implemented in the classroom, making it difficult for instructors to identify how to incorporate writing effectively. Drawing on the WTL literature, the MWrite program was established to work with STEM faculty to design, implement, and assess evidence-based WTL assignments. Herein we present a review of the WTL research generated through the MWrite program, situating our

findings in a four-dimensional framework of engagement to identify how the MWrite WTL assignment design and implementation has supported students' learning. Our analysis indicates that the multi-faceted design of MWrite WTL assignments supports students' development of conceptual knowledge and disciplinary thinking. The assignments' rhetorical features (i.e., context, audience, and genre) guide how students write about content, and peer review and revision stages encourage a collaborative, knowledge building process between students and their peers.

4.3 Introduction

Writing-to-learn (WTL) is an instructional practice that utilizes writing assignments to support students' learning. Investigations into the ways that writing can serve to develop knowledge were conducted as early as the 1970s.[1–4] However, research focused on the efficacy of writing assignments to support learning in STEM shows mixed results due in part to variation in how the assignments were implemented, what form they took, and the data gathered.[5,6] A series of research syntheses focused on how writing supports learning indicates that effective WTL assignments include elements that stimulate cognition and metacognition, provide meaning-making tasks, incorporate social interactions, and contain language that directs students towards specific learning goals.[7–10] The MWrite program at the University of Michigan was developed to support the uniform implementation of WTL assignments, across a variety of STEM courses, that incorporate the aforementioned elements of effective writing assignments while also minimizing barriers to implementation.[11–15] The goal of this article is to present a review of the research findings from courses involved in the MWrite program, providing insight into what forms of learning MWrite WTL assignment design can support and how the various design elements do so. The research findings are synthesized through the lens of an engagement framework.[16–18] to extend the findings beyond what is presented in the original research articles and to better understand how the MWrite WTL assignment design can holistically support student learning across the four dimensions of engagement (i.e., the cognitive, behavioral, emotional, and social dimensions).

4.3.1 WTL in STEM

Incorporating WTL in STEM can stimulate students' cognition while also appealing to affective and social elements of learning. The ability to appeal to multiple elements of learning indicates that WTL may support student engagement. While research has not yet directly addressed this potential connection between WTL pedagogy and student engagement, the WTL literature addresses elements that may be tied to the four dimensions of engagement. The existing literature describes how assignments engage students cognitively with both disciplinary concepts and disciplinary ways of thinking. A series of studies have detailed the benefits of writing assignments to improve students' scientific literacy or to elicit students' argumentation.[19–22] The Science Writing Heuristic (SWH) specifically supports developing students' conceptual knowledge, understanding of the nature of science, and disciplinary thinking during laboratory experiences by having them make explicit their observations, claims, and the evidence supporting their claims.[19,23–26] WTL research additionally provides insight into how students use multimodal representations in their writing and the types of learning WTL can support.[27–30] These studies on different implementations of writing in STEM courses illustrate how writing can stimulate students' cognition through clear, structured writing expectations.

Many of the WTL assignments described in the STEM education literature have students write in response to a particular audience or context.[20,30–33] Structuring WTL assignments such that students are writing to a specific audience can contribute to students' meaning-making while also simulating social interactions between them and the audience.[9,34] Studies describe assignments with a range of audiences, from the general public (often framed in the context of science communication; e.g., refs 30, 31) to more discipline-specific stakeholders (such as clients for some output; e.g., ref 32). Audiences can also include the teacher, students new to the content area, peers, or family members; research indicates that the audience influences what students write about and how.[35] Writing assignments can also become meaning-making tasks when they include a context relevant to students. Contextualizing scientific and mathematic content can support students in making inferences, evaluating content, and building connections to their lives.[33,36,37] For example, Balgopal and Wallace describe a style of WTL assignments implemented in biology courses where students apply their content knowledge to socioscientific issues, which are societally important issues that relate to the sciences.[20,38–41] They found that this style of assignment supported students' scientific literacy, argumentation, and use of abstract concepts. Altogether, these studies demonstrate how the audiences and contexts incorporated into WTL assignments can support students' learning. Viewed through the lens of engagement, the meaning-making supported by WTL may promote students' cognitive, affective, and social engagement as they consider content within a relevant context and describe content for a specific audience.

Social interactions have also been incorporated into WTL assignments by having students engage in peer review during the writing process. Calibrated Peer Review (CPR) has been incorporated into a range of introductory STEM courses, primarily for laboratory-oriented and disciplinary writing.[42] Various studies have found that students performed better on their writing and on questions associated with topics for which they engaged in CPR and demonstrated longer retention of the content in biology and chemistry courses;[42–46] furthermore, students often perceive the benefits of peer review associated with CPR.[47] The structured social interactions and students' perceptions of their value indicate that peer review may support both social and affective engagement with WTL.

The variety of research on WTL and peer review in STEM courses demonstrates the benefits of individual elements identified as important for effective WTL (i.e., stimulating students' cognition, creating meaning-making tasks, and incorporating social interactions). In addition, the elements of effective WTL align with task features thought to support student engagement (e.g., authentic tasks that provide students with opportunity for peer interactions).[16,48] However, to our knowledge, there is not yet research exploring the connections between the elements of effective WTL pedagogy and student engagement with learning. The WTL assignment design developed by the MWrite program differs from the WTL implementations described previously as it attempts to incorporate all of the reported elements of effective WTL; this makes the program an ideal space for exploring how the elements of effective WTL assignments may support learning across the four dimensions of engagement. The following section provides further background information on the MWrite program itself and how it served to support these principles across a variety of STEM courses.

4.3.2 Background on MWrite

MWrite was developed to support the implementation of WTL assignments with an evidence-based design while also attending to known barriers to faculty implementation of evidence-based practices. The MWrite program was created as part of an initiative at the University of Michigan to develop and support pedagogical innovations across the institution. Internal funding through the initiative, supplemented by two external grants, provided funding to support key personnel (i.e., instructors, writing fellows, and researchers) and develop a peer review tool to meet the needs of the program. MWrite is affiliated with the institution's center for writing, which additionally supports the program's institutionalization. Specifically, instructor and writing fellow training is conducted through the center for writing. Initially, five STEM instructors were recruited to participate in MWrite and further instructors have been recruited *via* word of mouth and symposia focused on pedagogical innovations within the institution. The MWrite program structure is outlined in Figure 4.1 and covered in more detail by Finkenstaedt-Quinn, Petterson, et al.[11]

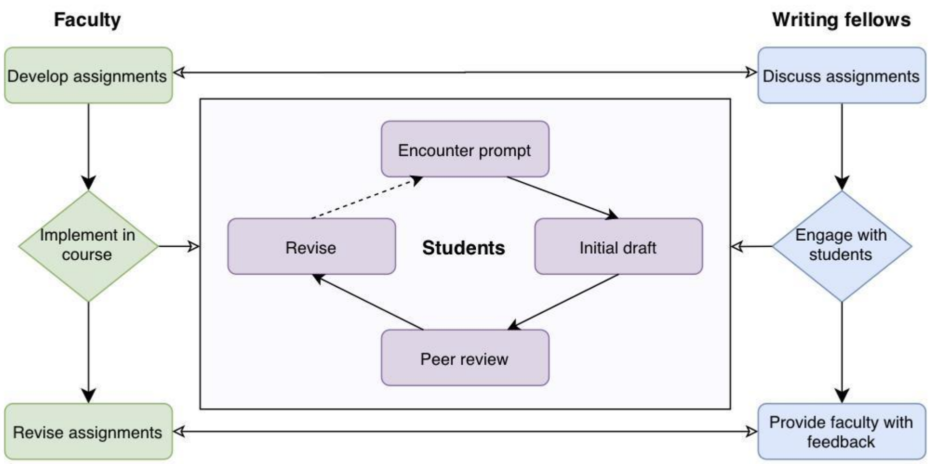


Figure 4.1. The MWrite Process. Students encounter prompts, write initial drafts, undergo peer review, and then submit revised drafts. Faculty design and implement the WTL assignments with support from writing fellows, who provide feedback on the assignments and interact with students as they respond to the assignments. Reprinted with permission from Finkenstaedt-Quinn, Petterson, Gere, and Shultz, *J. Chem. Educ.* 2021, 98, 5, 1548-1555. Copyright 2021 American Chemical Society.

The WTL assignment design is informed by the elements of effective WTL assignments.[7,9,10] Namely, assignments specify both a context and an audience, to create a meaning-making task and simulated social interactions. Additionally, all MWrite courses implement the WTL assignments in a process intended to encourage students' reflection and metacognition, whereby students submit initial drafts, participate in peer review, and submit revised drafts. The MWrite program provides instructors with support for both implementing WTL in their courses and designing WTL assignments. Faculty implementing MWrite are supported by access to writing fellows and an automated peer review tool. Writing fellows are undergraduate students with prior success in the course who provide current students with support on the MWrite assignments; they take a course through the university's writing center to train them in this work, meet regularly with faculty to discuss the assignments and content, and serve as a liaison between students and faculty in a way that allows for the improvement of MWrite assignments and the MWrite experience for future courses. The peer review tool interfaces with the university's online course management system and is designed to facilitate a double-blind peer review process in which students review the initial drafts and receive feedback from typically three peers. Both the writing fellows and automated peer review tool are intended to support the implementation of WTL in large-enrollment courses.

In addition to working with faculty to develop and implement WTL assignments, the MWrite program places a large emphasis on evaluating the effectiveness of the WTL assignment design and elucidating the modes by which the design supports students' learning. The goal for this review is to examine the outcomes of the assignment design approach of the WTL pedagogy implemented through the MWrite program, since various studies of WTL often implement different variations of the pedagogical approach.[5,6,8,9] The focus of this review specifically on MWrite research provides a summary and synthesis of the research findings coming out of this program with its standard implementation of WTL across a variety of STEM courses. Because this review is focused on the uniform design and implementation principles of WTL supported by the MWrite program, it can provide necessary insight on the types of learning instructors can support with the MWrite assignment design and how the various design components lead to and reinforce learning. This will contribute to the broader literature on WTL in STEM courses by providing insight into the ways WTL can support students' learning when implementing evidence-based assignment design principles. In addition, the synthesis of the research is guided by the framework of engagement to extend the understanding of how the MWrite assignment design supports learning across cognitive, behavioral, emotional, and social dimensions of learning.

4.4 Theory guiding this analysis

Writing and learning have been tied by a number of theories, generally aligned with three perspectives: writing as inherently supporting learning, writing as a cognitive process that supports learning, and writing as a sociocultural process that engages the learner.[9,49,50] The theories described within the literature as supporting WTL pedagogy are drawn from both cognitive and social theories of learning itself and of how writing can support learning. This is represented in our research as well, where the theories utilized in each MWrite research effort were dependent upon the research questions. The theories used range from purely cognitive (e.g., the cognitive process theory of writing; ref 51) to considering the social and contextual factors that influence learning and writing (e.g., the sociocultural theory of writing; ref 34). The goal of this review is to synthesize the findings of the MWrite research articles across learning domains, through the lens of *engagement*, to improve our understanding of student learning with WTL pedagogy and to inform the use of WTL beyond the MWrite context.

Engagement has been defined as a three-dimensional phenomenon that includes cognitive, behavioral, and emotional realms.[16–18] Cognitive engagement encompasses a student's psychological investment in learning and their strategic or self-regulated approach to learning, exemplified by persistence on challenging or difficult tasks, exerting effort to achieve mastery of ideas or skills, flexible approaches to problem solving, and use of metacognitive strategies, among others. Emotional engagement focuses on a student's affective domain, describing a student's feelings in an academic environment, such as interest, boredom, and anxiety. Lastly, behavioral engagement focuses on a student's behaviors in the classroom, such as effort, focus, and attention. More recently, some scholars have considered a social dimension as part of engagement, which incorporates interacting with peers in the academic context.[17,52] It is important to appeal to the multidimensional nature of engagement to create learning environments that can better support students' learning. For our purposes, we draw on the four-dimensional definition—with the cognitive, emotional, behavioral, and social dimensions—to characterize the ways in which the MWrite WTL assignments support students' engagement.

In the context of the MWrite WTL assignments, each of the four dimensions aligns with the ways in which students interact with the assignments. Cognitive engagement can be characterized by the conceptual understanding and disciplinary thinking that students demonstrate on the assignments, which serve as representations of students' persistence and effort to think critically about the content. Emotional engagement can be captured through students' reported attitudes and perceptions of the WTL assignments. Behavioral engagement can be thought of as the effort students demonstrate through completing the assignments and peer review, as well as students' self-reported effort. Lastly, social engagement is closely tied to behavioral engagement, where social engagement relates specifically to interactions with peers during the peer review process. We conceptualize social engagement as also extending to interactions with the writing fellows, interactions with peers beyond the structured review process, and choices made with respect to the audiences to whom students are writing. The four dimensions of engagement are interrelated, and individual students' experiences with the components of WTL pedagogy are likely mediated by different and multiple forms of engagement. However, by interpreting the findings of our review through the engagement framework lens, we can identify how the WTL assignments have elicited engagement across each of the four dimensions. This informs our understanding of how the WTL assignments may impact student engagement and how the elements of effective WTL contribute to supporting student learning across the four dimensions of engagement. In addition, focusing on engagement provides insight into the ability of the assignments to provide a holistic learning experience for students. Furthermore, interpretation of the findings from the MWrite program through the engagement framework lens is part of a systematic effort to evaluate the pedagogical impacts of utilizing WTL. In addition to interpreting the research findings from MWrite through the engagement framework lens, this review can serve as a platform to bring findings from discipline-based education research into conversation with the scholarship of teaching and learning. Specifically, the synthesis of research findings across courses, assignments, and research methodologies serves to demonstrate how WTL can support learning and engagement across contexts.

4.5 Methods

4.5.1 Reflexivity statement

This review synthesizes the research from the MWrite program at the University of Michigan. As this analysis is focused on research from MWrite specifically, it is important that we acknowledge our positionality as part of the MWrite program. Specifically, we are a program manager (SFQ), graduate student (FMW), and the co-primary investigators (GVS and ARG) working within the MWrite program. The co-first authors who engaged in the primary analysis for this article (SFQ and FMW) are co-authors on 12 of the 16 articles included in this analysis, and first authors on eight. This gave us familiarity with the research emerging from the MWrite program, which may have enhanced our ability to identify connections between the findings described in the articles. However, we recognize that our role in producing the research that was analyzed may have also led us to place a greater emphasis on minor findings of the articles that aligned with the themes identified in our analysis. In recognition of this potential for bias, we engaged in thematic analysis of the research articles with a consensus approach, as described in the analysis process section of the methods. In addition, prior to drafting this article, we discussed our findings with the entire MWrite team. Lastly, we sought feedback from researchers who were not affiliated with the MWrite program but are familiar with the research that has emerged from the program during the final drafting stages of this article.

4.5.2 Overview of articles included in the analysis

This review focuses on research articles about how students engaged with the WTL assignments implemented into classrooms using the MWrite WTL design. The analysis encompasses 16 research articles published from 2015 to 2022 (see 4.8 Appendix). The articles describe research on MWrite WTL assignments implemented in biology, chemistry, materials science, and statistics courses. The research broadly characterizes students' responses to the learning objectives of the assignments, gains in learning on those objectives, and students' experiences with the assignments. The data sources used in the articles include students' writing in response to components of the assignments, students' responses to pre/post external assessments of knowledge, student interviews, and students' responses to feedback surveys about the assignments. The qualitative data sources were analyzed through coding approaches and quantitative transformation; studies based on quantitative data typically used statistical analysis. The 4.8 Appendix presents the citation of each article and includes the disciplinary content area and a study overview for each article (Table 4.5).

4.5.3 Analysis process

Our review of the 16 research articles was guided by a qualitative thematic analysis approach.[53] The co-first authors (SFQ and FMW) separately read and wrote memos for each of the published MWrite WTL research articles with the intent to capture an overview of the study, the theoretical frameworks used, the data sources and methodologies, and the findings. Following this, we each read through our memos and identified themes in the findings across the articles. We then compared our themes and the articles we identified as contributing to each theme. The themes emerging from our independent analysis overlapped greatly, and through discussion we refined our themes into four categories, two focused on assessment of the learning objectives for the assignments (i.e., MWrite WTL assignments support students' abilities to describe content and leads to changes in content knowledge; MWrite WTL assignments engage students in disciplinary thinking practices) and two focused on how the structure of the WTL assignments support and scaffold learning (i.e., the structure of the MWrite prompts influences students' learning and affect; the peer review and revision processes support students' learning). For each theme, we additionally discussed how it aligned with the dimensions of engagement and gave insight into the mechanisms by which the WTL assignments supported learning. The final four thematic categories were discussed with the rest of the MWrite research team to confirm and finalize the analysis.

4.6 Results and discussion – MWrite research overview and analysis

4.6.1 Overview

Our thematic analysis of the research from the MWrite program indicates four categories drawn from the key claims and findings of the research articles. The four categories are that (1) MWrite supports students' abilities to describe content and that MWrite can lead to changes in students' content knowledge; (2) MWrite engages students in disciplinary thinking practices, specifically argumentation and reasoning; (3) the structure of MWrite assignments influences both students' learning and their affect towards the assignments; and (4) the peer review and revision process implemented with all MWrite assignments supports students' learning. The 4.8 Appendix provides an overview of which articles included in the analysis present findings pertaining to each category (Table 4.6). For each category, we synthesize the results of our analysis of the MWrite research articles and situate them in the engagement framework.

4.6.2 MWrite WTL assignments support students' abilities to describe content and lead to changes in content knowledge

A primary aim of the WTL assignments developed by the MWrite program is to support conceptual learning of STEM content in large introductory courses. Thus, much of the early research through the MWrite program focused on assessing whether this aim was achieved as students responded to the assignments. Eight articles describe research in this area, six of which each focused on a single writing assignment[54–59] and two focused on multiple writing assignments in a single course.[60,61] Across assignments, disciplines, and courses, there was evidence of students successfully describing content and demonstrating learning gains, as captured through the analysis of students' writing and/or revisions and through external assessments of students' knowledge (specific themes are presented in Table 4.1).

Table 4.1. Articles pertaining to each theme related to how MWrite WTL assignments support students' abilities to describe content and lead to changes in content knowledge

Article	Disciplinary content area	Themes				
		Analyses of student responses demonstrated students were able to describe content targeted by the WTL assignments	*The MWrite WTL design supports students' abilities to describe new or difficult concepts*	*The WTL assignments support students' application of content knowledge to real-world problems*	*Analysis of initial and revised drafts provides evidence of students' cognitive engagement*	*Students demonstrated learning via improvements on external assessments*
Finkenstaedt-Quinn et al., 2017 (ref 54)	Materials science and engineering	x			x	x
Moon, Zotos, et al., 2018 (ref 57)	Physical chemistry	x	x		x	x
Gere et al., 2018 (ref 60)	Introductory statistics	x		x		
Schmidt-McCormack et al., 2019 (ref 58)	Organic chemistry	x			x	x
Finkenstaedt-Quinn et al., 2020 (ref 56)	Physical chemistry	x		x	x	
Finkenstaedt-Quinn, Polakowski, et al., 2021 (ref 59)	Introductory statistics	x	x		x	
Brandfonbrener et al., 2022 (ref 55)	Organic chemistry	x	x			
Marks et al., 2022 (ref 61)	Materials science and engineering	x	x		x	x
Totals for each theme		8	4	2	6	4

Across articles, students demonstrated the ability to describe the content targeted by the WTL assignments (Table 4.1). In a majority of the articles, students' written responses to specific assignments were analyzed using rubrics that aligned with the learning objectives for the assignments.[54,56–61] For example, Finkenstaedt-Quinn et al. characterized students' responses to a WTL assignment implemented in an introductory materials science course focused on students' knowledge of material properties.[54] The assignment tasked students with extending their understanding of stress-strain properties of metals and ceramics to polymers, which are not discussed in as much detail in the course, and applying that knowledge to an authentic application. The analysis of students' writing indicated that students successfully extended their knowledge to the new material.[54] In Brandfonbrener et al., rather than characterizing students' responses using a rubric developed to align with the assignment objectives, the responses were characterized using an analytical framework developed from learning objectives for the fundamental chemistry concept of resonance.[55,62] The analysis indicated that students were incorporating descriptions of the concept in line with the established learning objectives. However, the results also indicated that some students' descriptions reflected surface-level conceptual understanding. This analysis also demonstrates the potential for WTL to elicit students' knowledge of difficult concepts when they are presented with challenging tasks, indicating students' cognitive engagement. Instructors can use the elicited knowledge to adapt their teaching of the material in an effort to move students from learning to mastery of the material.

Evidence indicating the potential of the MWrite WTL design to support students' learning is also found across MWrite studies. Specifically, studies indicate that the MWrite WTL design supports students' abilities to describe new or difficult concepts and that the assignments support students' application of content knowledge to real-world problems (Table 4.1). Supporting students in making connections between course content and real-world applications is a challenging learning goal to achieve in STEM education contexts.[63] Not only have we identified that students apply content knowledge to real-world problems in their writing, but students themselves recognize that the WTL assignments support them to build these connections.[64] This also demonstrates cognitive engagement with the assignments as students are applying effort to not only successfully describe new and difficult content in line with the learning objectives of the WTL assignments, but they are also applying content to real-world problems.

Further evidence of students' cognitive engagement was present in the articles in which both initial and revised drafts of students' writing were analyzed and compared (Table 4.1). The analyses demonstrated that students improved their descriptions of content upon revision. The improvement between drafts further demonstrates that the revision component of the assignments engaged students on the cognitive dimension, as revisions indicate students' persistence in the task and problem solving as they decide what feedback to incorporate as they revise. The improvement between the initial and revised drafts may also indicate students' behavioral engagement with the

WTL assignments, as it demonstrates that students made an effort to revise their drafts. Learning from responding to the assignments is also demonstrated by improvements seen on the external assessments in a subset of the studies (Table 4.1). In these articles, students who responded to the WTL assignments demonstrated greater gains on specific concepts targeted by the assignments than students in a control group who had not responded to the WTL assignments (but completed another activity related to the target concepts, such as a traditional problem set). The increase in students' conceptual knowledge through the MWrite WTL assignments aligns with findings from other implementations of WTL, such as the SWH.[24–26] Ultimately, the findings regarding students' learning gains on targeted course concepts provide evidence that the MWrite WTL design led to students' cognitive and behavioral engagement.

4.6.3 MWrite WTL assignments engage students in disciplinary thinking practices

Beyond engaging students in describing content and improving their content knowledge, another trend in the MWrite research is that the WTL assignments engage students in disciplinary thinking practices. Disciplinary thinking is a construct derived from the National Research Council's *A Framework for K-12 Education* that emphasizes the need for teaching in STEM classrooms to present science as a set of scientific and engineering practices, crosscutting concepts, and disciplinary core ideas.[65] The construct of disciplinary thinking relates to broader scientific and engineering practices, such as "analyzing and interpreting data," "constructing explanations," and "engaging in argument from evidence," using disciplinary concepts and ideas. While many MWrite research articles analyze students' construction of explanations within specific content areas, four MWrite research articles examined topics related to cognitively engaging students in other disciplinary thinking practices (Table 4.2).

Table 4.2. Articles pertaining to each theme related to how MWrite WTL assignments engage students in disciplinary thinking practices

Article	**Disciplinary content area**	**Themes**	
		The MWrite assignments support students' reflection on the nature of science	*The MWrite assignments support students' reasoning*
Shultz & Gere, 2015 (ref 66)	General chemistry	x	
Moon et al., 2019 (ref 67)	General chemistry		x
Watts et al., 2020 (ref 68)	Organic chemistry		x

Watts et al., 2022 (ref 69)	Organic chemistry		x
Totals for each theme		1	3

Analysis of students' writing and a pre/post survey in the study by Shultz and Gere indicate the capacity of the MWrite WTL assignments to support students in the scientific practices of asking questions and developing and using models, which is related to students' conceptions of the nature of science.[66] The scores of students' writing significantly increased from the initial to revised drafts for the learning objectives focused on describing and comparing scientific theories. However, students faced more challenges with comparing theories *versus* describing or summarizing a single theory. Results from a pre/post survey measure of students' conceptions of the nature of science indicated that students exhibited more sophisticated conceptions after the assignment, particularly for the idea that alternative theories in science exist. Hence, the study indicated that the MWrite WTL assignment was able to cognitively engage students in considering more deeply the scientific practices related to understanding the nature of science, in alignment with findings from the implementation of writing assignments through the SWH.[23]

The writing analysis in the other articles indicated that the MWrite assignments also support students' reasoning (Table 4.2), aligning with the scientific practices of constructing explanations and arguments. For example, the study by Moon et al. identified that students were able to make a variety of cognitive operations in their writing (e.g., observation, comparison, cause and effect) which could be characterized to determine the overall cognitive complexity in students' responses.[67] The study suggested that the measure of cognitive complexity is indicative of students' reasoning abilities. The two studies in organic chemistry focused on mechanistic reasoning, a specific type of scientific reasoning that requires explanation at a level lower than the observed phenomena.[70] Watts, Schmidt-McCormack, et al. analyzed features in students' writing necessary for this type of reasoning and identified that students were able to engage in multi-component, process-oriented reasoning, which is typically challenging for students.[68,71,72] Watts, Park et al. expanded on this work by examining how students reason on a meaning-making task reflective of the epistemic practices of organic chemists.[69] They found that the students who were more successful exhibited reasoning more aligned with how organic chemists would reason. In all three studies, the researchers separated correctness of content from identifying students' reasoning skills, suggesting that the WTL assignments are able to elicit both reasoning and content knowledge (as described in the previous section), in alignment with findings from implementations of the SWH.[19] This separation of content from reasoning skills further emphasizes how the WTL assignments can serve to cognitively engage students beyond developing their conceptual understanding and into engaging with disciplinary thinking, which can be a higher order task.

4.6.4 The structure of the MWrite assignments influences students' learning and affect

Various MWrite studies included specific findings about how the rhetorical features of the assignments—including the genres, contexts, and audiences—influence students' cognitive and social engagement with course content. The rhetorical framing of the assignments is a key aspect of MWrite WTL assignments, intended to support the writing activity as a meaning-making task. Studies across multiple course contexts demonstrated findings related to how the assignments' rhetorical features support students' learning and affect (Table 4.3).

Table 4.3. Articles pertaining to each theme related to how the structure of the MWrite assignments influences students' learning and affect

Article	**Disciplinary content area**	**Themes**			
		The audience or genre can influence the language students use and the degree to which students incorporate their content knowledge	*The assignment context supports students in making connections between concepts*	*Rhetorical framing is linked to students' perceptions of the relevance of content and their emotional engagement*	*Interview and survey data indicate that students often experience an increase in confidence due to WTL*
Finkenstaedt-Quinn et al., 2017 (ref 54)	Materials science and engineering	x			x
Moon, Zotos, et al., 2018 (ref 57)	Physical chemistry				x
Gere et al., 2018 (ref 60)	Introductory statistics	x			
Schmidt-McCormack et al., 2019 (ref 58)	Organic chemistry				x
Finkenstaedt-Quinn et al., 2020 (ref 56)	Physical chemistry		x		
Gupte et al., 2021 (ref 73)	Organic chemistry	x	x	x	
Petterson et al., 2022 (ref 64)	Organic chemistry	x	x	x	x

Marks et al., 2022 (ref 61)	Materials science and engineering		x	x	
Totals for each theme		4	4	3	4

Studies have indicated that the audience or genre can influence the language students use and the degree to which students incorporate their content knowledge (Table 4.3). For example, one study explored students' responses to two WTL assignments in a statistics course with different audiences and genres.[60] The findings indicated that the amount of statistics knowledge students incorporated differed between the two assignments, suggesting that the audience and genre can constrain how students incorporate their knowledge. Students recognize that the audience requires them to consider the detail of their explanations; this supports their perceived learning and can be a productive challenge for some students.[64,73] Thus, the audience and genre may influence the level of cognitive engagement with the course content, social engagement with the simulated audience, and behavioral engagement with the assignment itself.

Beyond the influence of the audience and genre on students' engagement, studies have demonstrated that the assignment context supports students in making connections between concepts targeted by the assignments, further eliciting students' cognitive engagement (Table 4.3). For example, two studies describe how the rhetorical context of specific WTL assignments supported students in making connections between concepts across microscopic and macroscopic scales,[56,61] which can be challenging connections for students to make.[74–76] Additionally, two studies indicated how the different contexts of WTL assignments supported students in making connections to content from prior courses, within the course itself, and within concurrently taken courses.[64,73] For example, interviews indicated that the assignments with medically relevant contexts supported students' perceived learning of the specific concepts targeted by the assignments.[64] The findings across the MWrite WTL implementation regarding the influence of the rhetorical context on student responses reflect similar findings from other implementations of WTL involving specific rhetorical contexts.[20,33,36,37,39–41] However, it is necessary to note that some students do not necessarily recognize the connections to outside courses, and that the context of the WTL assignments can influence whether students identify connections to content both within the course and from other sources, which can impact their cognitive engagement.[64,73]

The rhetorical framing is also linked to students' perceptions of the relevance of content and their emotional engagement, such as their affect, motivation, and confidence, with the assignments (Table 4.3). For example, interviews in an organic chemistry course revealed students' positive affective experiences with the role, genre, and audience assignment components because they contributed to the authenticity of the assignments and demonstrated the relevance of the content.[64] The relevance of the assignments supported students' motivation for learning and their identification of connections to fields of interest of future career possibilities. In addition, interview and survey data across studies have revealed that students often experience an increase in confidence due to WTL (Table 4.3). The increase in confidence has been further demonstrated when comparing to a control group engaged in a non-WTL, traditional homework activity[57] or when controlling for overall differences in academic ability.[58] The influence of the rhetorical framing on students' affect is similarly reported for other implementations of WTL.[33,36,37] However, studies within the MWrite context indicate that different students do have different affective experiences with assignment components.[61,64,73] For example, students may not always recognize the relevance of assignments, which can influence a negative affective experience for aspects of the WTL assignments that other students experience with positive affect.[73] Hence, attention within the WTL pedagogical approach must consider the fact that students will have different affective experiences that can influence their engagement.

4.6.5 The peer review and revision processes support students' learning

Peer review and revision are important stages of the MWrite WTL assignments that are intended to provide students with the opportunity to learn from their peers and revisit their own thinking. Our analysis indicates a few modes by which the two stages support students' learning and engagement. Ten of the MWrite WTL studies included some evaluation of the peer review and revision elements of the WTL assignments (Table 4.4). Examination of the findings across the studies indicates that the peer interactions occurring during the peer review process provide additional sources of knowledge for students that can lead to primarily content-focused revisions.

Table 4.4. Chapter 1 Articles pertaining to each theme related to how the peer review and revision processes support students' learning

Article	**Disciplinary content area**	**Themes**				
		Students can provide	*Students make*	*Students find reading*	*Reading peers' drafts*	*Peer review and revision*

		constructive, content-focused feedback to their peers, aligned with the peer review rubrics	*revisions associated with peer feedback they receive*	*peers' drafts useful for identifying whether they understood or explained content correctly*	*is statistically associated with content-focused revisions, exerting more influence than peer feedback received*	*reduced students' anxiety associated with the assignments and supported students' confidence*
Finkenstaedt-Quinn et al., 2017 (ref. 54)	Materials science and engineering			x		
Moon, Zotos, et al., 2018 (ref. 57)	Physical chemistry	x	x			
Halim et al., 2018 (ref. 77)	Introductory biology	x	x			
Finkenstaedt-Quinn et al., 2019 (ref. 78)	General chemistry	x	x			
Schmidt-McCormack et al., 2019 (ref. 58)	Organic chemistry		x			
Finkenstaedt-Quinn et al., 2020 (ref. 56)	Physical chemistry	x	x			
Gupte et al., 2021 (ref. 73)	Organic chemistry		x	x		x
Finkenstaedt-Quinn, Polakowski, et al., 2021 (ref. 59)	Introductory statistics				x	
Petterson et al., 2022 (ref. 64)	Organic chemistry		x	x		x
Watts et al., 2022 (ref. 69)	Organic chemistry		x		x	
Totals for each theme		4	8	3	2	2

Qualitative analyses of the peer review comments across multiple studies reveal that students can successfully provide constructive, content-focused feedback to their peers that aligns with the peer review rubrics students receive to guide the feedback process (Table 4.4). The alignment indicates that students are socially engaged during the peer review process, as well as cognitively engaged in the process of identifying content that merits revision and articulating feedback to their peers. Findings within Petterson et al. additionally indicate students' behavioral and social engagement with peer review, as interviewed students described putting more effort into their initial drafts so as to get the maximal benefit from peer feedback.[64] Some students also noted that if they produced a good first draft their peers could also benefit from reading their response.

A key theme related to peer review and revision was that peer feedback prompted students to make revisions on their drafts (Table 4.4). Relatedly, two studies characterized the features of peer review comments and revisions using logistic regression, finding that revisions were most associated with peer review comments that identified areas for improvement or that presented disagreements with their peers' reasoning.[69,78] These findings indicate cognitive engagement, as providing and utilizing feedback on content requires students to think about the material. Furthermore, analyses of student interviews and feedback surveys indicate that students describe peer feedback as helpful for identifying areas in their initial drafts that need improvement.[56,64,73] The studies also indicate social engagement during the revision process, as students are actively considering the feedback they received from their peers.

The importance of reading peers' drafts on students' revisions was also present in a subset of the studies. Analyses of student feedback survey responses and interviews indicated that students found reading peers' drafts to be beneficial for identifying whether they understood or explained content correctly (Table 4.4). Peers' drafts thereby serve as another source of knowledge, demonstrating an intersection between social and cognitive engagement with the WTL assignments. This finding is further supported by studies which used logistic regression to examine the relative influence of peer feedback and reading peers' drafts during the peer review process, which found that reading peers' drafts was statistically associated with content-focused revisions (Table 4.4). The impact of reading peers' drafts on students' own revisions indicates that students were socially and cognitively engaged while providing feedback to their peers. However, these studies suggested that peer feedback effected less influence on students' revisions relative to reading peers' drafts.[59,69]

A theme arising from more recent MWrite WTL research is the benefits of the peer review and revision stages of the assignments on student affect. Analysis of students' perceptions of the assignments indicate that both the peer review and revision stages of the MWrite process reduced students' anxiety associated with the assignments and supported students' confidence in their responses (Table 4.4). Students described how receiving feedback and reading their peers' responses made them feel more confident about their own responses. The positive affective responses to peer review align with the findings on students' perceptions of CPR as beneficial.[47]

In addition, Petterson et al. identified that the opportunity to revise can serve to reduce students' anxiety about the assignments more generally; students knew they could revise if they initially responded to the assignments incorrectly, which enabled them to take risks on their initial drafts.[64] In general, the inclusion of peer review and revision typically led to students' positive emotional engagement with the MWrite WTL assignments.

4.7 Conclusions and implications

This article analyzes the findings across the sixteen research articles extending from the MWrite program at the University of Michigan. The review encompasses impacts of the various aspects of the MWrite WTL implementation (i.e., assignment design, peer review, and revision) in a variety of introductory STEM disciplines, including chemistry, materials science, biology, and statistics. The findings from these articles were analyzed in alignment with the dimensions of cognitive, behavioral, emotional, and social engagement. Through analyzing the findings across articles from the MWrite program, this study extends the literature on WTL by focusing on students' engagement when WTL is uniformly implemented across STEM courses. That is, prior analyses and meta-analyses indicate that the design principles, implementation, and support structures of WTL assignments often varies across disciplines and courses. MWrite, however, provides instructional support for designing evidence-based, effective WTL assignments along with support for a peer review and revision process that is standard across MWrite courses. Hence, analyzing the set of findings extending from research on the MWrite program serves to provide insight into the ways that WTL, when implemented in classrooms following the principles behind MWrite, can support students' engagement in STEM.

We identified four key findings regarding how MWrite WTL supports students' engagement: (1) MWrite WTL assignments support students' abilities to describe content and leads to changes in content knowledge; (2) MWrite WTL assignments engage students in disciplinary thinking practices, specifically reasoning and argumentation; (3) the structure of MWrite assignments influences students' learning and affect; and (4) MWrite's peer review and revision processes support students' learning. The first two of these findings demonstrate the ways in which MWrite supports students' cognitive and behavioral engagement; the findings indicate that students are behaviorally engaged in the different aspects of WTL (i.e., drafting, peer review, and revising) which support their cognitive engagement with both content and disciplinary thinking practices. The third finding relates largely to students' emotional and social engagement, describing how the MWrite assignment design influences students' affective experiences and engagement with rhetorical contexts and audiences in ways that support their learning. The final finding relates to students' behavioral and social engagement, detailing how students participate in the peer review process which is grounded in the social aspects of receiving peer feedback and reading/responding to their peers' writing. Altogether, the findings illustrate how the multiple dimensions of engagement can be supported through MWrite WTL to create a more holistic learning experience for students.

The review points to a number of implications for instructors wishing to implement WTL in their courses and for stakeholders in programs seeking to support WTL across multiple courses and disciplines (such as writing across the curriculum or writing in the disciplines initiatives that seek to include WTL-specific support structures). The findings indicate that the design principles used for developing MWrite assignments created learning experiences that cognitively engaged students with both content and disciplinary thinking (such as scientific reasoning or arguing from evidence). This finding points to principles instructors should keep in mind when designing assignments, along with the benefits of implementing college- or university-wide programs like MWrite that can support students' learning through WTL assignments. In addition, the analysis indicates the potential for the WTL assignments to support learning in areas where students are known to struggle, such as building connections between concepts, connecting course content to real world applications, and engaging in complex reasoning. Furthermore, since WTL assignments such as those developed through MWrite elicit students' knowledge and disciplinary thinking, students' responses can serve as a valuable tool for formative assessment that can allow instructors to access what students know and understand about specific content areas. Lastly, the findings emphasize the value of implementing peer review and revision processes to support students' learning with WTL assignments. The various studies analyzed indicate the value of these processes for supporting students' learning, where both reading peers' writing and receiving feedback can inform content-focused revisions. Peer review and revision can additionally create positive affective experiences, such as increasing students' confidence.

The review additionally points to several avenues for further research into both WTL and the MWrite program specifically. Of particular need is research on whether and how the MWrite WTL assignment design may differentially impact groups of students and how the implementation and effectiveness of our WTL design changes in different classroom and institutional contexts. It is necessary to understand how students who are English language learners or who identify as belonging to minoritized groups experience the WTL assignments and the MWrite program. This direction for future research is especially merited given the findings that students express different experiences with different assignment components (such as the rhetorical contexts or peer review), both in terms of affect and in terms of their learning (e.g., the finding that not all students identify connections between assignment content and prior knowledge). Tied to research on the differential impact of WTL, more attention is required to understand the impacts of the WTL assignments on the affective domains, such as motivation and meaningful learning. For example, students may have different affective responses to the assignment components meant to demonstrate the relevance of course content (e.g., context and audience).

It is also important to study aspects of the MWrite program other than students' learning and engagement with the WTL assignments. Specifically, research focused on the impact of being involved with MWrite on both instructors and writing fellows is merited. As the MWrite program progresses, it is becoming increasingly apparent that there are unexpected benefits to the writing fellows and faculty involved (e.g., enculturation with disciplinary norms and increasing the use of evidence-based practices, respectively). Studying how being involved in a large-scale effort such as MWrite may influence pedagogy and disciplinary knowledge could inform our communities' efforts to create and support pedagogical change more broadly.

4.8 Appendix

Table 4.5. Overview of articles included in the review

<table>
<tr><td colspan="2">Shultz, G. V., & Gere, A. R. (2015). Writing-to-learn the nature of science in the context of the Lewis dot structure model. Journal of Chemical Education, 92(8), 1325-1329.</td></tr>
<tr><td>Disciplinary content area</td><td>General chemistry</td></tr>
<tr><td>Study overview</td><td>Researchers implemented a WTL assignment where students read and wrote about a fundamental article presenting a model for depicting structures in chemistry, comparing it to the conventional model taught in lecture. Students' responses to a pre-post survey indicated changes in their conceptions of the nature of science.</td></tr>
<tr><td colspan="2">Finkenstaedt-Quinn, S. A., Halim, A. S., Chambers, T. G., Moon, A., Goldman, R. S., Gere, A. R., & Shultz, G. V. (2017). Investigation of the Influence of a Writing-to-Learn Assignment on Student Understanding of Polymer Properties. Journal of Chemical Education, 94(11), 1610-1617.</td></tr>
</table>

Disciplinary content area	Materials science and engineering
Study overview	Researchers examined how a WTL assignment supported students' learning of polymer properties through analyzing students' writing with a content-focused rubric. The study additionally used a multi-tiered assessment and interviews to identify students' conceptual understanding and experiences with the writing process. The study found that WTL promoted students' understanding and identified aspects of WTL that supported their learning.
Moon, A., Zotos, E., Finkenstaedt-Quinn, S., Gere, A. R., & Shultz, G. (2018). Investigation of the role of writing-to-learn in promoting student understanding of light–matter interactions. *Chemistry Education Research and Practice, 19*(3), 807-818.	
Disciplinary content area	Physical chemistry
Study overview	Researchers investigated how students understand central concepts in physical chemistry through analyzing students' responses to a WTL assignment connecting the concepts to a real-world context. Using a quasi-experimental design and a pre-post assessment, the study identified learning gains associated with the WTL assignment. Findings were triangulated with interviews and feedback surveys and indicated that students improved their explanations of the concepts.
Halim, A. S., Finkenstaedt-Quinn, S. A., Olsen, L. J., Gere, A. R., & Shultz, G. V. (2018). Identifying and Remediating Student Misconceptions in Introductory Biology *Via* Writing-to-Learn Assignments and Peer Review. *CBE - Life Sciences, 17*(2), ar28.	
Disciplinary content area	Introductory biology
Study overview	Researchers examined four WTL assignments in an introductory biology course to identify the types of misconceptions elicited and how peer review and revision can remediate or propagate misconceptions. The study identified misconceptions in students' responses to all four assignments, and researchers generated six profiles to characterize how misconceptions were addressed through peer review. Findings indicated that directed peer review comments were the primary mode of remediating misconceptions, while students revealed further misconceptions when revising in response to more general peer review comments.
Gere, A. R., Knutson, A. V., Limlamai, N., McCarty, R., & Wilson, E. (2018). A Tale of Two Prompts: New Perspectives on Writing-to-Learn Assignments. *The WAC Journal, 29*, 147-188.	
Disciplinary content area	Introductory statistics
Study overview	Researchers analyzed students' responses to two WTL assignments in a statistics course, with a focus on how the amount and type of learning was influenced by the differences in genre and audience for the assignments. Responses were scored on a rubric to identify students' learning and interviews with students were conducted to identify the influence of genre and audience. The findings indicate that students' explanations differ based on the genre and the need to align the genre with the level of explanation targeted by the assignment.
Finkenstaedt-Quinn, S. A., Snyder-White, E. P., Connor, M. C., Gere, A. R., & Shultz, G. V. (2019). Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *Journal of Chemical Education, 96*(2), 227-237.	

<table>
<tr><td>Disciplinary content area</td><td>General chemistry</td></tr>
<tr><td>Study overview</td><td>Researchers investigated the relationships between peer review and revision through analyzing students’ peer review comments and revisions in their responses to the WTL assignment described in Shultz and Gere (2015; see above). Peer review comments were characterized by their usefulness and connected to associated revisions in students’ writing. The findings indicate that students provided detailed feedback that focused on concepts while also making editorial comments.</td></tr>
<tr><td colspan="2">Schmidt-McCormack, J. A., Judge, J. A., Spahr, K., Yang, E., Pugh, R., Karlin, A., Sattar, A., Thompson, B. C., Gere, A. G., Shultz, G. V. (2019). Analysis of the role of a writing-to-learn assignment in student understanding of organic acid–base concepts. Chemistry Education Research and Practice, 20(2), 383-398.</td></tr>
<tr><td>Disciplinary content area</td><td>Organic chemistry</td></tr>
<tr><td>Study overview</td><td>Researchers investigated a WTL assignment that required students to consider two theories of acid-base chemistry. The study included an external assessment administered to a treatment and comparison group, finding that students who completed the WTL assignment demonstrated a greater increase in their conceptual understanding. The results were triangulated with interviews and provide details about how students explained and connected the acid-base theories.</td></tr>
<tr><td colspan="2">Moon, A., Moeller, R., Gere, A. R., & Shultz, G. V. (2019). Application and testing of a framework for characterizing the quality of scientific reasoning in chemistry students' writing on ocean acidification. Chemistry Education Research and Practice, 20(3), 484-494.</td></tr>
<tr><td>Disciplinary content area</td><td>General chemistry</td></tr>
<tr><td>Study overview</td><td>Researchers investigated a WTL assignment focused on accessing students’ scientific reasoning. The study provides a framework for assessing students’ argumentative writing about ocean acidification, which was used to estimate the quality of students’ reasoning. The findings suggest strategies for identifying reasoning in students’ writing that can be used by instructors for formative assessment.</td></tr>
<tr><td colspan="2">Finkenstaedt-Quinn, S. A., Halim, A. S., Kasner, G., Wilhelm, C. A., Moon, A., Gere, A. R., & Shultz, G. V. (2020). Capturing student conceptions of thermodynamics and kinetics using writing. Chemistry Education Research and Practice, 21(3), 922-939.</td></tr>
<tr><td>Disciplinary content area</td><td>Physical chemistry</td></tr>
<tr><td>Study overview</td><td>Researchers identified students’ conceptions of two central concepts in physical chemistry through a WTL assignment that applied the concepts to a real-world context. The study focused on the content in students’ writing and the peer review feedback, finding that students demonstrated improvements in describing and connecting the concepts. The findings indicate that content-focused peer review and revision supported students’ responses to the assignment.</td></tr>
<tr><td colspan="2">Watts, F. M., Schmidt-McCormack, J. A., Wilhelm, C. A., Karlin, A., Sattar, A., Thompson, B. C., Gere, A. R., Shultz, G. V. (2020). What students write about when students write about mechanisms: analysis of features present in students’ written descriptions of an organic reaction mechanism. Chemistry Education Research and Practice, 21(4), 1148-1172.</td></tr>
<tr><td>Disciplinary content area</td><td>Organic chemistry</td></tr>
</table>

Study overview	Researchers analyzed features in students' writing in response to a WTL assignment about an organic chemistry reaction mechanism. The analysis adapted an analytical framework based in the philosophy of science to identify evidence of mechanistic reasoning in students' writing. Researchers analyzed the co-occurrences of features in students' writing to make inferences about students' reasoning, identifying empirical evidence for the hierarchical nature of mechanistic reasoning and the variations in students' reasoning.
Gupte, T., Watts, F. M., Schmidt-McCormack, J. A., Zaimi, I., Gere, A. R., & Shultz, G. V. (2021). Students' meaningful learning experiences from participating in organic chemistry writing-to-learn activities. *Chemistry Education Research and Practice*, *22*(2), 396-414.	
Disciplinary content area	Organic chemistry
Study overview	Researchers examined students' meaningful learning experiences from three WTL assignments in an organic chemistry laboratory course. The study analyzed students' responses to open-ended feedback surveys and interviews conducted after each assignment to understand if and how the assignments promoted students' meaningful learning across affective and cognitive domains. Findings indicated different ways the assignments connected to students' existing knowledge and the specific assignment components that supported students' meaningful learning.
Finkenstaedt-Quinn, S. A., Polakowski, N., Gunderson, B., Shultz, G. V., & Gere, A. R. (2021). Utilizing Peer Review and Revision to Support the Development of Conceptual Knowledge Through Writing. *Written Communication, 38*(3), 351-379.	
Disciplinary content area	Introductory statistics
Study overview	Researchers analyzed a WTL assignment, with a focus on identifying whether engaging in peer review and revision resulted in changes in how students write about the content elicited by the assignment. The findings demonstrate that students made content-focused revisions, including an increase in explaining content correctly. Furthermore, the study indicates that students benefit from reading peers' work during the peer review process.
Petterson, M. N., Finkenstaedt-Quinn, S. A., Gere, A. R., & Shultz, G. V. (2022). The Role of Authentic Contexts and Social Elements in Supporting Organic Chemistry Students' Interactions with Writing-to-Learn Assignments. *Chemistry Education Research and Practice*, *23*(1), 189-205.	
Disciplinary content area	Organic chemistry
Study overview	Researchers investigated WTL assignments in organic chemistry, with a focus on their inclusion of relevant contexts and social elements. Through analyzing interviews and feedback surveys, the study examined how the rhetorical elements of the WTL assignments demonstrated the relevance of organic chemistry and how peer review supported students' affective experiences. The findings indicated that assignments with relevance and social interactions support students' affective experiences and perceived learning.
Brandfonbrener, P. B., Watts, F. M., Shultz, G. V. (Accepted). Organic chemistry students' written descriptions and explanations of resonance and its influence on reactivity. *Journal of Chemical Education*, *98*(11), 3431-3441.	
Disciplinary content area	Organic chemistry

Study overview	Researchers examined students' responses to a WTL assignment focused on a concept in organic chemistry that is fundamental for representing and determining the reactivity of molecules. Through analyzing students' responses, the study identified how students explained the concept and how it influences reactivity. The analysis identified the features of the concept that students found important for their explanations, including the analogies and examples students generated. The findings indicated the ways students conceptualize the phenomenon.
Marks, L., Lu, H., Chambers, T., Finkenstaedt-Quinn, S., Goldman, R. S. (Accepted). Writing-to-learn in introductory materials science and engineering. *MRS Communications*, *12*, 1-11.	
Disciplinary content area	Materials science and engineering
Study overview	Researchers analyzed the influence of four WTL assignments on students' conceptual understanding for specific, targeted content areas. The researchers used scoring rubrics to analyze students' initial and revised drafts, finding statistically significant improvements in scores. The highest effect sizes were for the WTL assignments that required synthesizing qualitative data into quantitative formats. The researchers also used pre/post concept-inventory style assessments to identify that WTL supported students' learning beyond traditional pedagogies.
Watts, F. M., Park, G. Y., Petterson, M. P., Shultz, G. V. (2022). Considering alternative reaction mechanisms: Students' use of multiple representations to reason about mechanisms for a writing-to-learn assignment. *Chemistry Education Research and Practice*, *23*(2), 486-507.	
Disciplinary content area	Organic chemistry
Study overview	Researchers examined students' responses to a WTL assignment focused on how students utilized two representations fundamental in organic chemistry to determine and explain which pathway an organic chemistry reaction would follow. Through analyzing students' responses, the study identified how students explained their choice of reaction pathway and the changes in their explanations following revision. The analysis also identified the relative importance of the peer feedback students received and the peers' initial drafts that they read.

Table 4.6. Overview of articles pertaining to each category presented in the results and discussion

Article	**Themes**			
	MWrite WTL assignments support students' abilities to describe content and lead to changes in content knowledge	*MWrite WTL assignments engage students in disciplinary thinking practices*	*The structure of the MWrite assignments influences students' learning and affect*	*The peer review and revision processes support students' learning*
Shultz & Gere, 2015 (ref 66)		x		
Finkenstaedt-Quinn et al., 2017 (ref 54)	x		x	x
Moon, Zotos, et al., 2018 (ref 57)	x		x	x
Halim et al., 2018 (ref 77)				x
Gere et al., 2018 (ref 60)	x		x	

Finkenstaedt-Quinn et al., 2019 (ref 78)				x
Schmidt-McCormack et al., 2019 (ref 58)	x		x	x
Moon et al., 2019 (ref 67)		x		
Finkenstaedt-Quinn et al., 2020 (ref 56)	x		x	x
Watts et al., 2020 (ref 68)		x		
Gupte et al., 2021 (ref 73)			x	x
Finkenstaedt-Quinn, Polakowski, et al., 2021 (ref 59)	x			x
Petterson et al., 2022 (ref 64)			x	x
Brandfonbrener et al., 2022 (ref 55)	x			
Marks et al., 2022 (ref 61)	x		x	
Watts et al., 2022 (ref 69)		x		x
Total	8	4	8	10

4.9 Acknowledgements

We would like to acknowledge the University of Michigan Third Century Initiative, the Keck Foundation, the National Science Foundation (Grant No. 1524967), and the National Science Foundation Graduate Research Fellowship Program (Grant No. DGE 1256260) for funding. Additionally, we would like to thank all the faculty who have taken part in the MWrite Program and the students who have agreed to participate in our studies. Lastly, we would like to acknowledge Larissa Sano for her work developing the writing fellow and faculty seminars.

4.10 References

(1) Emig, J. Writing as a Mode of Learning. *Coll. Compos. Commun.* **1977**, *28* (2), 122–128.

(2) Galbraith, D. Conditions for Discovery through Writing. *Instr. Sci.* **1992**, *21* (1–3), 45–71. https://doi.org/10.1007/BF00119655.

(3) Bereiter, C.; Scardamlia, M. *The Psychology of Written Composition*; Lawrence Erbaum Associates, 1987.

(4) Flower, L.; Hayes, J. R. The Cognition of Discovery: Defining a Rhetorical Problem. *Coll. Compos. Commun.* **1980**, *31* (1), 21. https://doi.org/10.2307/356630.

(5) Arnold, K. M.; Umanath, S.; Thio, K.; Reilly, W. B.; McDaniel, M. A.; Marsh, E. J. Understanding the Cognitive Processes Involved in Writing to Learn. *J. Exp. Psychol. Appl.* **2017**, *23* (2), 115–127. https://doi.org/10.1037/xap0000119.

(6) Rivard, L. O. P. A Review of Writing to Learn in Science: Implications for Practice and Research. *J. Res. Sci. Teach.* **1994**, *31* (9), 969–983. https://doi.org/10.1002/tea.3660310910.

(7) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Res. Teach. English* **2015**, *50* (2), 199–235.

(8) Bangert-Drowns, R. L.; Hurley, M. M.; Wilkinson, B. The Effects of School-Based Writing-to-Learn Interventions on Academic Achievement: A Meta-Analysis Robert. *Rev. Educ. Res.* **2004**, *74* (1), 29–58.

(9) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and Conceptual Learning in Science: An Analysis of Assignments. *Writ. Commun.* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(10) Klein, P. D. Meddiators and Moderators in Individual and Collabborative Writing to Learn. *J. Writ. Res.* **2015**, *7* (1), 201–214.

(11) Finkenstaedt-Quinn, S. A.; Petterson, M.; Gere, A.; Shultz, G. Praxis of Writing-to-Learn: A Model for the Design and Propagation of Writing-to-Learn in STEM. *J. Chem. Educ.* **2021**, *98* (5), 1548–1555. https://doi.org/10.1021/acs.jchemed.0c01482.

(12) Moon, A.; Gere, A. R.; Shultz, G. V. Writing in the STEM Classroom: Faculty Conceptions of Writing and Its Role in the Undergraduate Classroom. *Sci. Educ.* **2018**, *102* (5), 1007–1028. https://doi.org/10.1002/sce.21454.

(13) Trafimow, D.; Ruckel, L.; Stovall, S.; Raut, Y. Predicting Faculty Intentions to Assign Writing in Their Classes. *Int. J. Scholarsh. Teach. Learn.* **2017**, *11* (2).

(14) Stroumbakis, K. D.; Moh, N.; Kokkinos, D. Community College STEM Faculty Views on the Value of Writing Assignments. *WAC J.* **2016**, *27* (1), 142–154. https://doi.org/10.37514/wac-j.2016.27.1.08.

(15) Finkenstaedt-Quinn, S. A.; Gere, A. R.; Dowd, J. E.; Thompson Jr., R. J.; Halim, A. S.; Reynolds, J. A.; Schiff, L. A.; Flash, P.; Shultz, G. V. Postsecondary Faculty Attitudes and Beliefs About Writing-Based Pedagogies in the STEM Classroom. *CBE Life Sci. Educ.* **2022**, Accepted.

(16) Fredricks, J. A.; Blumenfeld, P. C.; Paris, A. H. School Engagement: Potential of the Concept, State of the Evidence. *Rev. Educ. Res.* **2004**, *74* (1), 59–109. https://doi.org/10.3102/00346543074001059.

(17) Finn, J. D.; Zimmer, K. S. Student Engagement: What Is It? Why Does It Matter? In *Handbook of Research on Student Engagement*; Christenson, S. L., Wylie, C., Reschly, A. L., Eds.; 2012; pp 97–131. https://doi.org/10.1007/978-1-4614-2018-7.

(18) Fredricks, J. A.; Wang, M.-T.; Linn, J. S.; Hofkens, T. L.; Sung, H.; Parr, A.; Allerton, J. Using Qualitative Methods to Develop a Survey Measure of Math and Science Engagement. *Learn* **2016**, *43*, 5–15. https://doi.org/http://dx.doi.org/10.1016/j.learninstruc.2016.01.009.

(19) Grimberg, B. I.; Hand, B. Cognitive Pathways: Analysis of Students' Written Texts for Science Understanding. *Int. J. Sci. Educ.* **2009**, *31* (4), 503–521. https://doi.org/10.1080/09500690701704805.

(20) Balgopal, M. M.; Wallace, A. Writing-to-Learn, Writing-to Communicate, & Scientific Literacy. *Am. Biol. Teach.* **2013**, *75* (3), 170–175. https://doi.org/10.1525/abt.2013.75.3.5.

(21) Klein, P. D. Constructing Scientific Explanations through Writing. *Instr. Sci.* **2004**, *32* (3), 191–231. https://doi.org/10.1023/B:TRUC.0000024189.74263.bd.

(22) McDermott, M. A.; Hand, B. A Secondary Reanalysis of Student Perceptions of Non-Traditional Writing Tasks over a Ten Year Period. *J. Res. Sci. Teach.* **2010**, *47* (5), 518–539. https://doi.org/10.1002/tea.20350.

(23) Keys, C. W.; Hand, B.; Prain, V.; Collins, S. Using the Science Writing Heuristic as a Tool for Learning from Laboratory Investigations in Secondary Science. *J. Res. Sci. Teach.* **1999**, *36* (10), 1065–1084. https://doi.org/10.1002/(SICI)1098-2736(199912)36:10<1065::AID-TEA2>3.0.CO;2-I.

(24) Hand, B.; Wallace, C. W.; Yang, E. M. Using a Science Writing Heuristic to Enhance Learning Outcomes from Laboratory Activities in Seventh-Grade Science: Quantitative and Qualitative Aspects. *Int. J. Sci. Educ.* **2004**, *26* (2), 131–149. https://doi.org/10.1080/0950069032000070252.

(25) Hand, B.; Hohenshell, L.; Prain, V. Examining the Effect of Multiple Writing Tasks on Year 10 Biology Students' Understandings of Cell and Molecular Biology Concepts. *Instr. Sci.* **2007**, *35* (4), 343–373. https://doi.org/10.1007/s11251-006-9012-3.

(26) Poock, J. R.; Burke, K. A.; Greenbowe, T. J.; Hand, B. M. Using the Science Writing Heuristic in the General Chemistry Laboratory to Improve Students' Academic Performance. *J. Chem. Educ.* **2007**, *84* (12), 1371–1379. https://doi.org/10.1021/ed084p2007.

(27) Hand, B.; Choi, A. Examining the Impact of Student Use of Multiple Modal Representations in Constructing Arguments in Organic Chemistry Laboratory Classes. *Res. Sci. Educ.* **2010**, *40* (1), 29–44. https://doi.org/10.1007/s11165-009-9155-8.

(28) McDermott, M. A.; Hand, B. The Impact of Embedding Multiple Modes of Representation within Writing Tasks on High School Students' Chemistry Understanding. *Instr. Sci.* **2013**, *41* (1), 217–246. https://doi.org/10.1007/s11251-012-9225-6.

(29) Gunel, M.; Kingir, S.; Aydemir, N. The Effect of Embedding Multimodal Representation in Non-Traditional Writing Task on Students' Learning in Electrochemistry. In *Using Multimodal Representations to Support Learning in the Science Classroom*; Hand, B., McDermott, M., Prain, V., Eds.; 2016; pp 59–75.

(30) McDermott, M. A.; Hand, B. Using Multimodal Representations to Support Learning in the Science Classroom. In *Using Multimodal Representations to Support Learning in the Science Classroom*; Hand, B., McDermott, M., Prain, V., Eds.; 2016; pp 183–211. https://doi.org/10.1007/978-3-319-16450-2.

(31) Rootman-Le Grange, I.; Retief, L. Action Research: Integrating Chemistry and Scientific Communication to Foster Cumulative Knowledge Building and Scientific Communication Skills. *J. Chem. Educ.* **2018**, *95* (8), 1284–1290. https://doi.org/10.1021/acs.jchemed.7b00958.

(32) Herrington, A. J. Writing in Academic Settings: A Study of the Contexts for Writing in Two College Chemical Engineering Courses. *Res. Teach. English* **1985**, *19* (4), 331–361. https://doi.org/10.4324/9781003059219-11.

(33) Doe, S.; Pilgrim, M. E.; Gehrtz, J. Stories and Explanations in the Introductory Calculus Classroom: A Study of WTL as a Teaching and Learning Intervention. *WAC J.* **2016**, *27* (1), 94–118. https://doi.org/10.37514/wac-j.2016.27.1.06.

(34) Prior, P. A Sociocultural Theory of Writing. In *Handbook of Writing Research*; MacArthur, C., Graham, S., Fitzgerald, J., Eds.; Guilford, 2006; pp 54–66.

(35) Gunel, M.; Hand, B.; McDermott, M. A. Writing for Different Audiences: Effects on High-School Students' Conceptual Understanding of Biology. *Learn. Instr.* **2009**, *19* (4), 354–367. https://doi.org/10.1016/j.learninstruc.2008.07.001.

(36) Rathburn, M. K. Building Connections Through Contextualized Learning in an Undergraduate Course on Scientific and Mathematical Literacy. *Int. J. Scholarsh. Teach. Learn.* **2015**, *9* (1). https://doi.org/10.20429/ijsotl.2015.090111.

(37) Libarkin, J.; Ording, G. The Utility Ofwriting Assignments in Undergraduate Bioscience. *CBE Life Sci. Educ.* **2012**, *11* (1), 39–46. https://doi.org/10.1187/cbe.11-07-0058.

(38) Balgopal, M. M.; Casper, A. M. A.; Wallace, A. M.; Laybourn, P. J.; Brisch, E. Writing Matters: Writing-to-Learn Activities Increase Undergraduate Performance in Cell Biology. *Bioscience* **2018**, *68* (6), 445–454. https://doi.org/10.1093/biosci/biy042.

(39) Balgopal, M. M.; Montplaisir, L. M. Meaning Making: What Reflective Essays Reveal

about Biology Students' Conceptions about Natural Selection. *Instr. Sci.* **2011**, *39* (2), 137–169. https://doi.org/10.1007/s11251-009-9120-y.

(40) Balgopal, M. M.; Wallace, A. M.; Dahlberg, S. Writing to Learn Ecology: A Study of Three Populations of College Students. *Environ. Educ. Res.* **2012**, *18* (1), 67–90. https://doi.org/10.1080/13504622.2011.576316.

(41) Balgopal, M. M.; Wallace, A. M.; Dahlberg, S. Writing from Different Cultural Contexts: How College Students Frame an Environmental SSI through Written Arguments. *J. Res. Sci. Teach.* **2017**, *54* (2), 195–218. https://doi.org/10.1002/tea.21342.

(42) Russell, A. A. The Evolution of Calibrated Peer Review^{TM}. *ACS Symp. Ser.* **2013**, *1145*, 129–143. https://doi.org/10.1021/bk-2013-1145.ch009.

(43) Cox, C. T.; Poehlmann, J. S.; Ortega, C.; Lopez, J. C. Using Writing Assignments as an Intervention to Strengthen Acid-Base Skills. *J. Chem. Educ.* **2018**, *95* (8), 1276–1283. https://doi.org/10.1021/acs.jchemed.8b00018.

(44) Mynlieff, M.; Manogaran, A. L.; Maurice, M. S.; Eddinger, T. J. Writing Assignments with a Metacognitive Component Enhance Learning in a Large Introductory Biology Course. *CBE Life Sci. Educ.* **2014**, *13* (2), 311–321. https://doi.org/10.1187/cbe.13-05-0097.

(45) Pelaez, N. J. Problem-Based Writing with Peer Review Improves Academic Performance in Physiology. *Am. J. Physiol. - Adv. Physiol. Educ.* **2002**, *26* (1–4), 174–184. https://doi.org/10.1152/advan.00041.2001.

(46) Gunersel, A. B.; Simpson, N. Improvement in Writing and Reviewing Skills with Calibrated Peer ReviewTM. *Int. J. Scholarsh. Teach. Learn.* **2009**, *3* (2). https://doi.org/10.20429/ijsotl.2009.030215.

(47) Ruggiero, D.; harbor, jon. Using Writing Assignments with Calibrated Peer Review to Increase Engagement and Improve Learning in an Undergraduate Environmental Science Course. *Int. J. Scholarsh. Teach. Learn.* **2013**, *7* (2). https://doi.org/10.20429/ijsotl.2013.070221.

(48) Newmann, F. M.; Wehlage, G. G.; Lamborn, S. D. The Significance and Sources of Student Engagement. In *Student Engagement and Achievement in American Secondary Schools*; Newmann, F. M., Ed.; Teachers College Press: New York, NY, 1992; pp 11–39. https://doi.org/10.4324/9780203012543-16.

(49) Klein, P. D.; Boscolo, P. Trends in Research on Writing as a Learning Activity. *J. Writ. Res.* **2016**, *7* (3), 311–351. https://doi.org/10.17239/jowr-2016.07.03.01.

(50) Klein, P. D.; Arcon, N.; Baker, S. Writing to Learn. In *Handbook of Writing Research*; MacArthur, C. A., Graham, S., Fitzgerald, J., Eds.; The Guilford Press, 2015.

(51) Flower, L.; Hayes, J. R. A Cognitive Process Theory of Writing. *Coll. Compos. Commun.* **1981**, *32* (4), 365–387.

(52) Linnenbrink-Garcia, L.; Rogat, T. K.; Koskey, K. L. K. Affect and Engagement during Small Group Instruction. *Contemp. Educ. Psychol.* **2011**, *36* (1), 13–24. https://doi.org/10.1016/j.cedpsych.2010.09.001.

(53) Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101.

(54) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Chambers, T. G.; Moon, A.; Goldman, R. S.; Gere, A. R.; Shultz, G. V. Investigation of the Influence of a Writing-To-Learn Assignment on Student Understanding of Polymer Properties. *J. Chem. Educ.* **2017**, *94* (11), 1610–1617. https://doi.org/10.1021/acs.jchemed.7b00363.

(55) Brandfonbrener, P. B.; Watts, F. M.; Shultz, G. V. Organic Chemistry Students' Written Descriptions and Explanations of Resonance and Its Influence on Reactivity. *J. Chem. Educ.* **2021**, *98* (11), 3431–3441. https://doi.org/10.1021/acs.jchemed.1c00660.

(56) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Kasner, G.; Wilhelm, C. A.; Moon, A.; Gere, A. R.; Shultz, G. V. Capturing Student Conceptions of Thermodynamics and Kinetics Using Writing. *Chem. Educ. Res. Pract.* **2020**, *21* (4), 922–939. https://doi.org/10.1039/C9RP00292H.

(57) Moon, A.; Zotos, E.; Finkenstaedt-Quinn, S.; Gere, A. R.; Shultz, G. Investigation of the Role of Writing-to-Learn in Promoting Student Understanding of Light-Matter Interactions. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 807–818. https://doi.org/10.1039/c8rp00090e.

(58) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(59) Finkenstaedt-Quinn, S. A.; Polakowski, N.; Gunderson, B.; Shultz, G. V.; Gere, A. R. Utilizing Peer Review and Revision in STEM to Support the Development of Conceptual Knowledge Through Writing. *Writ. Commun.* **2021**. https://doi.org/10.1177/07410883211006038.

(60) Gere, A. R.; Knutson, A. V.; Limlamai, N.; McCarty, R.; Wilson, E. A Tale of Two Prompts: New Perspectives on Writing-to-Learn Assignments. *WAC J.* **2018**, *29* (1), 147–188. https://doi.org/10.37514/wac-j.2018.29.1.07.

(61) Marks, L.; Lu, H.; Chambers, T.; Finkenstaedt-Quinn, S.; Goldman, R. S. Writing-to-Learn in Introductory Materials Science and Engineering. *MRS Commun.* **2022**, *12*, 1–11. https://doi.org/10.1557/s43579-021-00114-z.

(62) 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.

(63) Gilbert, J. On the Nature of “Context” in Chemical Education. *Int. J. Sci. Educ.* **2006**, *28* (9), 957–976. https://doi.org/10.1080/09500690600702470.

(64) Petterson, M. N.; Finkenstaedt-Quinn, S. A.; Gere, A. R.; Shultz, G. V. The Role of Authentic Contexts and Social Elements in Supporting Organic Chemistry Students’ Interactions with Writing-to-Learn Assignments. *Chem. Educ. Res. Pract.* **2022**, *23* (1), 189–205. https://doi.org/10.1039/d1rp00181g.

(65) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; The National Academies Press, 2012.

(66) Shultz, G. V.; Gere, A. R. Writing-to-Learn the Nature of Science in the Context of the Lewis Dot Structure Model. *J. Chem. Educ.* **2015**, *92* (8), 1325–1329. https://doi.org/10.1021/acs.jchemed.5b00064.

(67) Moon, A.; Moeller, R.; Gere, A. R.; Shultz, G. V. Application and Testing of a Framework for Characterizing the Quality of Scientific Reasoning in Chemistry Students’ Writing on Ocean Acidification. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 484–494. https://doi.org/10.1039/c9rp00005d.

(68) Watts, F. M.; Schmidt-McCormack, J.; Wilhelm, C.; Karlin, A.; Sattar, A.; Thompson, B.; Gere, A. R.; Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students’ Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21*, 1148–1172. https://doi.org/10.1039/C9RP00185A.

(69) Watts, F. M.; Park, G. Y.; Petterson, M. N.; Shultz, G. V. Considering Alternative Reaction Mechanisms: Students’ Use of Multiple Representations to Reason about Mechanisms for a Writing-to-Learn Assignment. *Chem. Educ. Res. Pract.* **2022**, *23* (2), 486–507. https://doi.org/10.1039/d1rp00301a.

(70) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Sci. Educ.* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(71) Bhattacharyya, G.; Bodner, G. M. “It Gets Me to the Product”: How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402–1407. https://doi.org/10.1021/ed082p1402.

(72) Bhattacharyya, G. Who Am I? What Am I Doing Here? Professional Identity and the Epistemic Development of Organic Chemists. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 84–92. https://doi.org/10.1039/b806222f.

(73) Gupte, T.; Watts, F. M.; Schmidt-McCormack, J. A.; Zaimi, I.; Gere, A. R.; Shultz, G. V. Students' Meaningful Learning Experiences from Participating in Organic Chemistry Writing-to-Learn Activities. *Chem. Educ. Res. Pract.* **2021**, *22* (2), 396–414. https://doi.org/10.1039/d0rp00266f.

(74) Bain, K.; Moon, A.; Mack, M. R.; Towns, M. H. A Review of Research on the Teaching and Learning of Thermodynamics at the University Level. *Chem. Educ. Res. Pract.* **2014**, *15* (3), 320–335. https://doi.org/10.1039/c4rp00011k.

(75) Bain, K.; Towns, M. H. A Review of Research on the Teaching and Learning of Chemical Kinetics. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 246–262. https://doi.org/10.1039/c5rp00176e.

(76) Justi, R. Teaching and Learning Chemical Kinetics. In *Chemical Education: Towards Research-based Practice*; Gilber, J. K., de Jong, O., Justi, R., Treagust, D. F., van Driel, J. H., Eds.; 2002; pp 293–315.

(77) Halim, A. S.; Finkenstaedt-Quinn, S. A.; Olsen, L. J.; Gere, A. R.; Shultz, G. V. Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review. *CBE Life Sci. Educ.* **2018**, *17* (2), 1–12. https://doi.org/10.1187/cbe.17-10-0212.

(78) Finkenstaedt-Quinn, S. A.; Snyder-White, E. P.; Connor, M. C.; Gere, A. R.; Shultz, G. V. Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *J. Chem. Educ.* **2019**, *96* (2), 227–237. https://doi.org/10.1021/acs.jchemed.8b00711.

Chapter 5
Investigating Writing-To-Learn To Support Organic Chemistry Students' Meaningful Learning Experiences

5.1 Initial remarks

This chapter examines how writing-to-learn (WTL) assignments in organic chemistry specifically can support students' learning. Building on the MWrite literature synthesized in Chapter 4, this chapter focuses specifically on the implementation of WTL in the second-semester organic chemistry laboratory course. Through examining students' feedback surveys and interviews corresponding to their experiences with the WTL assignments administered in the course, this study highlights how the assignments supported students' meaningful learning experiences. The study provides the motivation for focusing on WTL as the primary instructional practice for the dissertation, as the findings indicate that WTL can both support and elicit students' understanding of course concepts while also supporting students' motivation and interest.

The study specifically uses meaningful learning theory to interpret student responses to feedback surveys and interviews for each of the three WTL assignments in the second-semester organic chemistry laboratory course. Meaningful learning theory emphasizes the interplay between cognitive, affective, and psychomotor domains within learning experiences, and focuses on the constructivist definition of learning as connecting new knowledge to prior knowledge. The study explores the connection between the cognitive and affective domains as promoted through the WTL assignments. Results from the study highlight the various ways that WTL assignments can support students in making connections between new concepts and their existing knowledge, specifically through building on their prior knowledge, extending their understanding of course concepts, and making connections to concepts from other chemistry courses often taken concurrently. For example, the first WTL assignment in the course focused on acid–base concepts, and students' responses indicated that they found the assignment useful for helping them build connections between their understanding of acids and bases from prior courses, so they could apply their knowledge of acid–base concepts to their learning in the organic chemistry laboratory.

Additionally, the results highlight how various features of the WTL assignments promote the affective components of meaningful learning. Specifically, students reported how the assignments required them to engage in solving challenging problems, how the audience and genre of the assignments promoted their interest and motivation, and how the peer review process improved their confidence. Key implications from this research include details for how instructors can construct WTL assignments or similar tasks by selecting topics which can build on students' existing knowledge and through incorporating assignment design components that can support students' interest and motivation—such as specifying an audience or genre which might relate to their future career interests.

This chapter was originally published as a research article in *Chemistry Education Research and Practice*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author, I contributed to conceptualization, methodology, data analysis, and writing (both original draft preparation and review and editing). I shared primary authorship with T. Gupte, an undergraduate mentee, who contributed to conceptualization and data analysis. J.A. Schmidt-McCormack contributed to conceptualization, methodology, data collection, analysis, and writing (original draft preparation). I. Zaimi contributed to data analysis. A.R. Gere and G.V. Shultz contributed to funding acquisition, project supervision, conceptualization, data collection, and writing (review and editing).

<u>*Original publication and copyright information:*</u>

5.2 Abstract

Teaching organic chemistry requires supporting learning strategies that meaningfully engage students with the challenging concepts and advanced problem-solving skills needed to be successful. Such meaningful learning experiences should encourage students to actively choose to incorporate new concepts into their existing knowledge frameworks by appealing to the cognitive, affective, and psychomotor domains of learning. This study provides a qualitative analysis of

students' meaningful learning experiences after completing three writing-to-learn (WTL) assignments in an organic chemistry laboratory course. The assignments were designed to appeal to the three domains necessary for a meaningful learning experience, and this research seeks to understand if and how the WTL assignments promoted students' meaningful learning. The primary data collected were the students' responses to open-ended feedback surveys conducted after each assignment. These responses were qualitatively analyzed to identify themes across students' experiences about their meaningful learning. The feedback survey analysis was triangulated with interviews conducted after each assignment. The results identify how the assignments connected to students' existing knowledge from other courses and indicate that assignment components such as authentic contexts, clear expectations, and peer review supported students' meaningful learning experiences. These results inform how assignment design can influence students' learning experiences and suggest implications for how to support students' meaningful learning of organic chemistry through writing.

5.3 Introduction

Teaching and learning in organic chemistry are challenging because the discipline is highly conceptual and requires advanced problem-solving and critical thinking skills. These challenges are compounded by the need for students to develop specific learning strategies that may not directly transfer from general to organic chemistry.[1–3] Because it is a challenging course even for students who are successful in general chemistry, organic chemistry classrooms can be high-stress environments with high rates of attrition.[1,4–6] In response, chemistry education researchers are invested in understanding and measuring students' meaningful learning experiences in organic chemistry.[7–10] Theories of meaningful learning address the challenges with learning organic chemistry by recognizing the interplay between affective, cognitive, and psychomotor components of learning. Meaningful learning theories posit that all three of these areas must be addressed for meaningful learning to occur.[11–13] To further support students' learning in organic chemistry, it is necessary to develop and research specific pedagogical approaches to support students' meaningful learning.

5.3.1 Meaningful learning in organic chemistry

Studies of meaningful learning in organic chemistry have theoretical grounding in Ausubel's theory of meaningful learning and Novak's theory of human constructivism.[11,13,14]

These frameworks both draw on the constructivist definition of learning as connecting new knowledge to prior knowledge. Ausubel posits three requirements for meaningful learning: (1) students' relevant prior knowledge, (2) instructors' organization of concepts to relate new information to students' prior knowledge, and (3) students actively choosing to incorporate new knowledge into their existing conceptual frameworks.[14] Of these three requirements, only the second is within the instructors' control. Hence, it is necessary that instructors' curricular choices support students in relating new concepts to their prior knowledge. However, instructors can indirectly influence the third of these requirements by developing curricular materials that build sufficient interest in new concepts, encouraging students to actively make connections to their prior knowledge. Ausubel's theory of meaningful learning is related to Novak's theory of human constructivism, which encompasses three domains related to learning: the cognitive, affective, and psychomotor domains.[11,13] The cognitive domain includes engaging conceptual knowledge and reasoning skills, the affective domain relates to attitudes and motivation towards learning, and the psychomotor domain involves motor skills and physical movement. Each of these domains is required for meaningful learning; that is, a learning experience must attend to all three domains for it to be meaningful.

Research on meaningful learning in the organic chemistry curriculum focuses on both faculty and student perspectives. Bretz et al. interviewed chemistry faculty about their goals for teaching undergraduate lab courses across general, organic, and upper-division laboratory courses.[15] They analyzed faculty responses through the lens of meaningful learning frameworks and found that faculty held goals pertaining to all three domains of meaningful learning. However, their analysis suggested that faculty teaching organic chemistry exhibited a marked decrease in their discussion of affective goals, instead emphasizing critical thinking and laboratory techniques. The affective goals that instructors did have included emphasizing the relevance of the content and skills taught in the laboratory to students' aspirations. This study suggests that organic chemistry instructors need to incorporate learning activities that specifically support the affective domain, because many laboratory courses already focus on the cognitive and psychomotor domains.

Meaningful learning has also been characterized from the students' perspective. Galloway and Bretz developed the Meaningful Learning in the Laboratory Instrument (MLLI), which measures students' meaningful learning experiences in laboratory courses across the cognitive and affective domains.[7] Administrations of the MLLI to students across general and organic chemistry suggest that students tend to have a variety of cognitive and affective expectations and experiences in chemistry laboratories.[8,9] Galloway and Bretz' national, cross-sectional study indicated that students with low affective expectations tended to have experiences that fulfilled their negative expectations.[8] Further, their longitudinal study indicated that students' experiences can fail to meet their positive expectations. However, students' expectations tended to reset when beginning organic chemistry lab courses, even when expectations for general chemistry labs were not met.[9] That students' expectations are reset before organic chemistry provides opportunities to re-engage students with the affective domain of meaningful learning in this course. As each of these studies indicates, pedagogy in organic chemistry would benefit from attention to the affective dimension of meaningful learning.

Research qualitatively exploring chemistry students' affective experiences as they relate to laboratory learning experiences is limited.[16] A study by Galloway et al. focused on the interplay between students' affective, cognitive, and psychomotor experiences by interviewing students in both general and organic chemistry laboratories.[10] They found that the affective domain is closely linked to the cognitive and psychomotor domains and that students' affective experiences are linked to their approach to learning. In particular, their analysis suggested that students' differing sense of autonomy in the laboratory influenced a wide range of approaches to learning, from rote to meaningful strategies. Their findings indicate the need for further qualitative research exploring students' affective experiences. Furthermore, their results indicate a need for the design of learning experiences which help students develop a sense of autonomy and specifically support students' positive affective experiences. This finding is particularly important when considering the recent attention on developing hybrid and online laboratory courses.[17,18] Research has suggested that general chemistry students completing virtual laboratory experiences develop similar cognitive and psychomotor skills as students completing in-person laboratories but report lower affective experiences.[17,18] These studies, in particular, identify the importance for both hybrid and online laboratories to emphasize the value of laboratory learning experiences to students' lives and their career aspirations. Hensen et al. specifically call for learning interventions that can allow students to have positive affectual laboratory experiences.[18]

5.3.2 Writing-to-learn and meaningful learning

The existing research on students' meaningful learning experiences in organic chemistry laboratories suggests the need to explore pedagogical approaches that can support students' positive affective experiences. Such affective experiences relate to students' interest and motivation. These constructs are often aligned with relevancy, which, in turn, is influenced by instructors' curricular choices.[19] Prior research within chemistry indicated that lesson plans which appeal to topics relevant to students' lives can improve motivation for learning.[20] Studies of motivation for learning more generally have provided evidence for various influences that teachers' pedagogical choices can have on students' motivation.[21–23] Specifically, setting challenging goals for students, explaining the rationale for assignments, supporting learning from peer models, and providing timely feedback have been shown to support students' motivation.[21,22] In addition, students can have higher-quality motivation for when they perceive a teaching environment as providing clear structure and supporting their autonomy.[23]

Prior research within chemistry education has investigated appealing to students' interest and motivation by developing curricula that incorporate authentic contexts.[24,25] A component of these efforts requires the design of specific course materials that are relevant to students, which is important for encouraging students to make connections between new concepts and existing knowledge structures.[2] Therefore, it is necessary to research assignments within the organic chemistry context that are specifically designed to appeal to the affective domain of meaningful learning by supporting students' interests.

Writing-to-learn (WTL) activities are instructional interventions that can complement the psychomotor domain emphasized in the laboratory by specifically appealing to the cognitive and affective learning domains. They are designed to engage students with material while supporting their conceptual understanding through the process of writing, with a focus on improving content knowledge rather than improving writing ability.[26–28] Research on WTL assignments has been conducted in chemistry,[29–35] biology,[36] and engineering courses,[37] demonstrating how WTL assignments with peer review and revision serve to elicit students' content knowledge while supporting students' understanding of targeted concepts.

While WTL assignments have been demonstrated to support students' conceptual learning, research is necessary to investigate the components of WTL assignments that engage students in meaningful learning. Previous literature has identified that WTL prompt design is important for supporting students' conceptual learning.[38,39] Notably, one of the essential components is "meaning-making," defined as the requirement that WTL assignments have students apply their existing knowledge to new situations. Other important features are that WTL assignments have rhetorical components to highlight the relevance of content to authentic situations—such as a specified genre, role, and audience—and that WTL assignments include interactive components such as peer review.[38,39] While these studies have illustrated assignment components that support students' conceptual learning, it is valuable to also understand students' meaningful learning experiences with WTL assignments. This understanding is necessary, because students—rather than instructors—are responsible for making the decisions to integrate new ideas into their existing knowledge structures.

Research on undergraduate students' meaningful writing experiences has found that assignments that engage students with both content and peers while connecting to students' current and future identities are more meaningful for students.[40] To build upon these findings, it is necessary to specifically research students' meaningful learning experiences with WTL assignments in STEM courses. Furthermore, prior studies of WTL interventions in STEM courses have focused on evaluating the implementation of a single assignment rather than examining the outcomes of implementing a series of WTL assignments within a course. Analyzing a series of WTL assignments is valuable, especially for understanding similarities and differences between individual assignments and how, when implemented throughout a course, they might support students' meaningful learning experiences over a semester.

5.4 Research questions

This research presents the qualitative analysis of second-semester organic chemistry laboratory students' meaningful learning experiences with a set of writing-to-learn assignments. We describe each of the three WTL assignments administered to students, and our study is focused on thematic analysis of students' responses to feedback surveys that elicited their cognitive and affective experiences with the assignments. Two research questions specifically guide this study:

1. How do organic chemistry students experience building connections between new concepts and their existing knowledge when responding to writing-to-learn assignments?
2. What components of writing-to-learn assignments do students perceive as supporting their learning of organic chemistry course content?

5.5 Theoretical framework

This research is guided by the aforementioned theories of meaningful learning, with additional attention to developing relevancy through authentic tasks and considering students' motivation for learning.[20–23,41] We posit that WTL assignments meet Ausubel's three requirements for meaningful learning while appealing to the learning domains in Novak's theory of human constructivism.[11,13,14] As with any learning experience, the ability for WTL assignments to appeal to the elements of these learning theories is dependent on each assignment, the instructional context, and students' previous learning experiences. Nevertheless, the WTL assignments central to this study were designed with specific components meant to support students' meaningful learning. WTL assignments, in general, can be designed such that they require students to use their previous knowledge while exploring new concepts.[38,39] For example, an organic chemistry WTL assignment has the potential to help students transfer their existing knowledge of acid–base chemistry to their learning of the electron-pushing formalism. Furthermore, prior research has shown that WTL assignments can encourage students to connect new concepts to their prior learning.[37] Hence, we suggest that WTL assignments can be designed to provide the opportunity for students to choose to connect new information to their prior knowledge, thereby appealing to the cognitive learning domain.

In addition to appealing to the cognitive domain, WTL assignments can appeal to the affective learning domain by including rhetorical components meant to interest students by presenting authentic situations in which the content is relevant. Relevancy in science education, as described by Stuckey et al., contains three dimensions: the individual, societal, and vocational.[19] Within WTL assignments, the connection of target concepts to specific contexts has the ability to appeal to one or more of these, dependent upon the individual learner and the context within the assignment. For example, an organic chemistry WTL assignment within the context of medicinal chemistry has the possibility to appeal to the societal domain by addressing the impacts of introducing new pharmaceuticals into society. Such an assignment could additionally appeal to the vocational domain for students interested in practicing medicine. The use of contexts within WTL assignments can thereby engage with both personal and societal dimensions, thereby moving beyond incorporating a context as a simple association between a single concept and a specific application. In this way, WTL assignments have the opportunity to appeal to the model of context-based curricula that focuses on social circumstances, which is theorized to most effectively incorporate context into conceptual learning.[24] We posit that, through carefully selected rhetorical components, WTL assignments have the opportunity to appeal to the affective learning domain.

Relevancy and authentic tasks are closely related to students' motivation, which also relates to the affective learning domain. This is of particular importance, as students—and not instructors—are responsible for choosing to incorporate new knowledge into their existing knowledge framework. Prior research has used self-determination theory to characterize motivation for learning as rooted in motives that are either autonomous (i.e., by choice) or controlled (i.e., not by choice).[22] Studies have shown that students' motives—and the quality of their motivation—can be influenced by the teaching environment, such as the language used for assignments and instructions, the timeliness of feedback, and providing clear rationales for learning activities.[21–23] Each of these elements of the teaching environment are important considerations when implementing WTL assignments and are therefore valuable when considering students' experiences of WTL assignments. By interpreting students' experiences from completing the WTL assignments through theoretical perspectives of meaningful learning, relevancy through authentic tasks, and students' motivation, we can identify if and how the WTL assignments encourage students to engage in meaningful learning. Furthermore, these frameworks allow for insight into the specific assignment components and implementation structures that might support students' learning.

5.6 Methods

To address our research questions, we employed a qualitative design that allowed for a rich understanding of the range of students' meaningful learning experiences with the WTL assignments.[42] The qualitative design, in particular, complements the existing quantitative research on students' meaningful learning in organic chemistry.[7–9,16] To broadly examine students' meaningful learning experiences, the primary data source for this research is second-semester organic chemistry laboratory students' responses to open-ended feedback survey questions that were administered after the completion of each of the three WTL assignments. Semi-structured interviews conducted after each WTL activity served as a secondary data source to triangulate and corroborate the findings that emerged from analyzing the feedback survey responses.

5.6.1 Setting and participants

This study took place at a large Midwestern research university in the January–April 2018 semester. Three WTL assignments were administered in the second-semester organic chemistry laboratory course, which included a weekly one-hour lecture and four-hour laboratory period. At this institution, the laboratory course is offered separately from the second-semester lecture course. The lecture course is a prerequisite/corequisite for the laboratory course. Historically, 84% of students take the lecture and laboratory courses simultaneously. In addition to the WTL assignments, students completed laboratory reports and took quizzes for assessment. The three WTL assignments contributed approximately 20% towards students' grades for the course. The participants for this study include the students enrolled in the course (N = 695), specifically those who opted to respond to optional feedback surveys (N = 333, 149, and 147, respectively for each assignment) and participate in interviews (N = 10, 9, and 8, respectively for each assignment). All students who participated in the surveys and interviews provided their informed consent, and Institutional Review Board permission was granted for this study.

5.6.2 Writing-to-learn assignment design and implementation

The WTL assignments were designed and implemented with attention to the four essential characteristics for successful assignments as identified in Gere et al.'s review of WTL prompts: (1) engaging students in applying content-knowledge to a new task, (2) incorporating structures for peer interactions during the writing process, (3) supporting metacognition and reflection by requiring revision, and (4) setting clear expectations for what students should include in their writing.[39] Each WTL assignment targeted different content areas to engage students' application of knowledge to new situations.

The first WTL assignment focused on acid–base chemistry. The prompt identified levothyroxine, a drug for treating hypothyroidism, and discussed how its effectiveness differs when interacting with different calcium supplements. Students were to assume the role of a medicinal chemist and to write an email to a physician with whom they were collaborating on a study about the co-administration of levothyroxine with calcium supplements. The objectives for the assignment were for students to describe how the levothyroxine molecule could act as a sodium salt, how a calcium ion could act as a Lewis acid, and why one calcium supplement would inhibit the absorption of levothyroxine but a different calcium supplement would not. These objectives require understanding the relationship between pH and pK_a and understanding how pH affects molecules' protonation states. The focus on acid–base chemistry was meant to reinforce the concepts students were formally introduced to in their prior organic chemistry experiences.

The second WTL assignment focused on a modified, base-free Wittig reaction and its substrate scope. The prompt described a base-free catalytic Wittig reaction and the implications for performing the reaction on an industrial scale in the production of chemicals such as vitamin A. Students were to assume to role of a science reporter for Chemical and Engineering News, and to write an article describing how the base-free Wittig mechanism related to the standard Wittig mechanism. In the article, students were also to discuss how the new reaction required no base and to discuss the limitations in substrate scope as identified within the prompt. The objectives for the assignment were for students to describe the traditional Wittig reaction, to determine the mechanism for a base-free modification of the Wittig reaction, and to discuss the substrate limitations for the base-free Wittig reaction. While students learned the traditional Wittig reaction during the lecture component of the laboratory course, this assignment was meant to encourage students to extend that knowledge by considering an alternative reaction that avoided the use of an external base.

The third and final WTL assignment focused on the reactivity of the drug thalidomide. This prompt described the history of the drug thalidomide being used as a sedative with harmful side effects. The assignment identified racemization and acid hydrolysis as mechanisms that affect thalidomide, and placed students in the role of an organic chemist writing a grant proposal about thalidomide analogues that would prevent these mechanisms. The objectives for this assignment were for students to describe the racemization and amide acid hydrolysis mechanisms for the thalidomide molecule, to propose an analogue that would prevent these mechanisms, and to explain how NMR could be used to monitor the progress of the hydrolysis reaction. This assignment was intended to relate broadly to the knowledge students should have been exposed to across their experiences in organic chemistry, including the concurrently taken second-semester lecture course. Specifically, the assignment was meant to reinforce the general mechanisms for racemization and acyl transfer reactions, both of which are formally taught in the second-semester lecture course.

The three assignments will hereafter be referred to as the acid–base, Wittig, and thalidomide assignments, respectively. The full text of each assignment is provided in 5.11.1 Appendix 1. For each WTL assignment, students had one week to write their first draft, four days to complete the peer review process, and four days to revise and submit a second draft. Structures for peer interactions were provided by an automated peer review process in which each student read drafts and provided feedback to typically three peers. During the weeks the assignment components were open, students had the opportunity for further peer interaction with the course writing fellows. The writing fellows were undergraduate students who were previously successful in the course and trained to assist students with the content of the WTL assignments. The revision assignment after the peer review process provided time for students to revise their assignment after reflecting on their initial draft, peers' drafts they had read, and feedback they had received. Initial drafts and peer review comments were assessed for completion, and students were provided a rubric indicating the content areas that would be the focus of assessment for their final drafts. The assessment process was independent of the research reported herein.

5.6.3 Data collection

After the students turned in the second draft of each writing assignment, they were provided a link to a feedback survey. The survey asked students to respond to the following questions:

1. What do you like about this assignment? Please describe any aspects of the presentation or content of this writing assignment that were unclear.
2. What did you find the most challenging to write about?

Responding to the feedback surveys was not required, and students were not offered points toward their final course grade or other incentives for completing the feedback surveys. Of the 695 students enrolled in the course, 333 (48%) responded to the acid–base assignment feedback survey, 149 (21%) responded to the Wittig assignment feedback survey, and 147 (21%) responded to the thalidomide assignment feedback survey. All survey responses were included in the analysis. Additionally, semi-structured interviews were conducted after students had completed all three components of each assignment (N = 10, 9, and 8 for the acid–base, Wittig, and thalidomide assignments, respectively).[43] Students were recruited to participate in the interviews through a question at the end of the feedback surveys. These interviews were conducted as part of a larger research effort to understand students' responses and experiences with the WTL assignments. The interview protocol included some portions related to students' learning experiences, including questions such as "What did you learn by doing this assignment?" and "Is there anything that you found challenging to write about?" Each interview was audio recorded with the students' permission.

5.6.4 Data analysis

The feedback survey responses were qualitatively analyzed with a coding scheme developed to characterize the students' meaningful learning experiences across all three WTL assignments. Codes were inductively developed for each question of the feedback survey for a single assignment. While researchers recursively coded the feedback survey questions for all three assignments, the coding scheme continued to be modified and developed.[42] After the initial development of the coding scheme, two researchers (TG and JSM) coded the same subset of responses and discussed their application of codes for each response (N = 60; 20 surveys from each assignment; 9.5% of the total surveys). Modifications were made to clarify the coding scheme, and a consensus was reached for the codes applied to these responses.

Two researchers (TG and JSM) then independently coded 20% of the feedback surveys not used in the development of the coding scheme and met to discuss the application of codes. Fuzzy kappa, a modified version of Cohen's kappa that allows for multiple codes to be applied to a single response, was calculated to determine the reliability of the coding scheme.[44] A fuzzy kappa value of 0.82 was calculated for the coding of students' responses to the first survey question, and a value of 0.85 was calculated for the coding of the second survey question. These values both indicate strong agreement among the two researchers.[45] Any disagreements for applying the coding scheme to individual responses were then resolved to reach a consensus for the final application of codes. One researcher (TG) then coded the remaining feedback survey responses.

The finalized coding scheme that was applied to all responses included two broad categories corresponding to students' responses to the two survey questions. The first category captures students' positive and negative affective experiences with the assignments, whereas the second category captures the challenges students had with the assignments. Each of these categories contains codes representing different aspects of students' meaningful learning experiences as elicited by the two feedback survey questions, and multiple codes could be applied to each response. After coding, the research team organized codes across both categories of the coding scheme into specific themes. The thematic analysis and the corresponding codes are presented in the Results. The frequencies of codes across survey responses to each assignment are presented alongside the complete coding scheme in 5.11.2 Appendix 2, Table 5.3.

The interview data were used to corroborate findings related to the different meaningful learning experiences reported across the feedback surveys. As the interviews were used as a secondary data source to triangulate the primary feedback survey data, portions of the interviews related to the themes emerging in the feedback survey analysis were considered. All interviews were transcribed verbatim, then reviewed by the research team to identify excerpts of students describing their meaningful learning experiences. The research team then met to discuss the excerpts identified across all interviews and identify connections to the feedback survey analysis.

5.7 Results and discussion

The goal of this research is to characterize organic chemistry students' meaningful learning experiences with WTL assignments. To do this, we administered feedback surveys and conducted interviews to understand students' perceptions of the WTL assignments. Our analysis sought to understand how the assignments encouraged students to build connections between new concepts and existing knowledge. Furthermore, our analysis focused on the components of the WTL assignments and implementation that served to support students' meaningful learning with specific attention to the cognitive and affective learning domains. The results are organized by our two research questions. Each section is supplemented with excerpts from the feedback survey responses, while excerpts from the interview responses that corroborate each theme are provided in 5.11.3 Appendix 3, Table 5.4.

5.7.1 How do organic chemistry students experience building connections between new concepts and their existing knowledge when responding to writing-to-learn assignments?

Students reported that the WTL assignments encouraged them to make connections between new concepts and existing knowledge in different ways. This finding appeared in students' responses to both survey questions across the three assignments. Each WTL assignment appeared to support students' perceptions of how they built connections to existing knowledge in slightly different ways: *via* application of knowledge from previous courses, from the current course, and from a concurrent course. We captured these connections through the overarching theme building connections between content, and we have summarized the key sub-themes for this research question in Table 5.1. We will first describe the findings broadly relating to this theme, then we examine each assignment individually to illustrate the different ways students identified the assignments led them to connect new content to their existing knowledge.

Table 5.1. Sub-themes related to RQ1: How do organic chemistry students perceive building connections between new concepts and their existing knowledge when responding to the writing-to-learn assignments?

Sub-theme	Exemplar
Building on prior knowledge	*Acid-base feedback survey:* "I liked that this assignment helped reinforce the concepts learned in [Organic Chemistry 1] about acids and bases and their overall effect in a chemical reaction."
Building on course concepts	*Wittig feedback survey:* "What I liked about this writing assignment was that it pertained to the type of reaction that we had done in class."
Building on concepts from a concurrent course	*Thalidomide feedback survey:* "I liked how well it tied into what we were learning in [Organic Chemistry II Lecture]. It made it easy to understand why the mechanism proceeded in the way that it did."

Building connections between content. This theme encompasses instances of students describing how the assignments served to connect the new content presented within each WTL assignment to their existing knowledge. The most common responses were instances in which students described being challenged by writing about new concepts introduced by each assignment. Students finding the newly introduced concepts to be challenging provides evidence that the WTL prompts met their intended objectives of encouraging students to engage with new and challenging concepts through the writing process. Hence, the writing assignments appeared to successfully appeal to the cognitive domain required for a meaningful learning experience while providing tasks sufficiently challenging, which can support students' academic motivation.[22] The feedback survey responses in which students indicated challenges with the conceptual material were closely examined to gain further insight for each prompt. From examining these responses, we found that the problems posed by each of the WTL assignments related to students' existing knowledge in slightly different ways: the acid–base assignment by connecting to students' knowledge from previous courses; the Wittig assignment by building upon students' knowledge gained in the laboratory course itself; and the thalidomide assignment by building on students' knowledge gained from the concurrent second-semester organic chemistry lecture course. These details demonstrate the range of ways in which writing assignments can engage students in applying their existing knowledge to new topics in a meaning-making task.[38,39] Each of the ways the assignments connected to students' prior knowledge are described in more detail below.

The remaining findings within this theme capture students' affective experiences regarding how the assignments helped them build connections between content. Some students indicated general positive feelings about how the assignments focused on developing their conceptual understanding. Other students indicated appreciating how they were required to draw from prior knowledge to address questions posed by the assignments. Students expressed mixed affective experiences about how closely aligned they found the assignments to the course. Each of these findings relate to how each assignment appealed to the affective domain of the learning experience and supported students in building connections to different types of existing knowledge. That students referred to the cognitive, conceptual components of the assignments with affective language contributes to prior research findings which suggested the inherent relationships between the cognitive and affective domains of meaningful learning.[10]

Acid–base assignment: building on prior knowledge. Students' feedback on the acid–base WTL assignment revealed that they perceived the need to know and apply their knowledge from previous chemistry courses. Students reported being challenged by each learning objective of the assignment, aligning with results from our previous work investigating the concepts that posed challenges for students on the same WTL prompt that was administered in the previous year.[34] Many of the acid–base concepts that students reported challenges with, including difficulties with the relationship between pH and pK_a and defining or applying definitions of Lewis acids have been previously reported in the literature as challenging topics for students.[34,35,46–51]

Results from the feedback survey and interview analyses indicated that students perceived drawing upon their prior knowledge to complete the acid–base assignment. This use of prior knowledge is evident from instances in which students indicated, with positive affect, the need to use knowledge gained from other courses to formulate an answer to the assignment. For example, one student wrote,

> "I liked that it was a review of the things we have learned in [Organic Chemistry I Lecture], and we had to put different topics together in order to really answer the questions being asked."

However, some students did not necessarily recognize that recall of prior knowledge was necessary for completing the assignment, finding the content irrelevant to the laboratory course: "It was also frustrating because it seemed that there were gaps between what has been taught and what we were supposed to know."

The prevalence of students using prior knowledge to respond to the acid–base assignment was identified in the interview responses (e.g., the excerpt from Gabriella's interview in 5.11.3 Appendix 3, Table 5.4). Similarly, some students not recognizing the connections to prior knowledge was also evident in the interviews (e.g., the excerpt from Matthew's interview presented in 5.11.3 Appendix 3, Table 5.4). These responses suggest that some students found the assignment difficult because the content within the acid–base assignment was not directly related to the content they were currently learning. While this provides further evidence that the acid–base assignment required integration of previously acquired knowledge about acid–base chemistry, it also suggests that these conceptual connections between the organic chemistry courses were not explicitly clear to all students. This finding suggests that it is important to not only implement assignments that encourage meaningful learning by requiring students to connect to prior knowledge, but also to explicitly clarify the underlying and fundamental concepts. This is especially necessary for concepts from general chemistry that prior research on faculty perceptions suggests are important for students' success in organic chemistry.[52]

Wittig assignment: building on course concepts. The analysis of the feedback surveys after the second WTL assignment revealed that students found it to be challenging because of the way it extended ideas from the laboratory course itself. The concepts students primarily described as being challenging were related to two of the learning goals: how the modified Wittig reaction could function without a base and why acrylate would not participate in the modified Wittig reaction when the structurally similar maleate would. These challenges were reflected in the feedback survey responses; for example, one student wrote about needing to consider how the base-free Wittig reaction was both similar and different from the standard Wittig reaction the students performed in the laboratory:

> "I enjoyed thinking more about the reaction we learned about in class from a different angle. It was interesting to think about using no base, as well as various schemes that were similar to what we did in lab."

Similar responses appeared in the interviews, in which students further expressed being challenged by the assignment's requirement to reflect upon and develop an account for why the base-free Wittig reaction works in some cases but not others (e.g., the excerpt from Jameson's interview in 5.11.3 Appendix 3, Table 5.4). These challenges are reflective of the conceptual development expected to be elicited by case-comparison problems in organic chemistry.[53] This finding contributes to the literature by identifying that students experience the challenges intended by such problems.

The components of the assignment for which students indicated positive affect reflect how the assignment related to topics from the lecture and lab components of the course, where they were introduced to and performed the standard Wittig reaction. For example, one feedback response indicated:

> "I like that this assignment exposed us to things going on in the chemistry world today that are tied to the reaction we did in lab. I feel like we did not necessarily learn all the information we needed to answer the question posed in the writing assignment."

However, as this response suggests, not all students valued that the assignment required them to extend their knowledge of the reaction performed in the laboratory to new situations. Despite this, other students indicated favoring the structure of the assignment, which first asked students about more familiar material (the traditional Wittig reaction) before asking students questions about new material (the modified Wittig reaction). For example, some students explicitly mentioned the way the assignment connected new material to what they were already familiar with:

> "I enjoyed the challenging aspect of the intramolecular mechanism present in the assignment. This made me apply what I already know to a new concept I was not too familiar with."

This response aligns with prior research that suggests students are able to apply concepts they are familiar with to new material when responding to WTL assignments.[37] Furthermore, this finding contributes to our understanding of WTL assignments by identifying that some students recognize and value WTL assignments that require them to extend their existing knowledge. Overall, students' feedback on the Wittig assignment suggests how instructors can organize assignments to help students build connections between concepts in such a way that encourages them to recognize and choose to integrate new knowledge into their existing conceptual understanding.

Thalidomide assignment: building on concepts from a concurrent course. Similar to the Wittig assignment, students also struggled with the new concepts that were presented by the thalidomide assignment. In the feedback survey analysis, students mentioned being challenged by each of the assignment's learning goals. For example, one student wrote about being challenged to determine the acid hydrolysis mechanism:

"I found it rather difficult to come up with a mechanism for the acid hydrolysis products. It was difficult in that a student could come up with multiple mechanisms that worked, but had no particular hint as to why one mechanism may be favored over another."

Students' abilities to describe the acid hydrolysis reaction mechanism are explained in our prior work, which demonstrates how some students, but not all, were able to connect explanatory concepts to the steps in the mechanism.[35] Additionally, students were challenged by the requirement to explain how the mechanism could be monitored by NMR. One student wrote:

"For me the most challenging part was figuring out how to use NMR to determine the reaction progress. After realizing that it was the peaks that mattered, it made much more sense."

This difficulty reflects prior research demonstrating students' challenges when reasoning about proton NMR spectra.[54] Our results indicate that students' perceived challenges for this assignment align with the intended challenges for the assignment, and extend the literature related to these concepts by identifying that students do perceive the inherent challenges.

Students' challenges with the assignment are reflective of the fact that the content addressed draws from concepts across the introductory organic chemistry curriculum. For instance, the acid hydrolysis mechanism was taught to students in the concurrent second-semester organic chemistry lecture course, and many students indicated this connection with positive affect. This is exemplified by a feedback survey response in which a student wrote: "I like that it made me use what I have learned in [Organic Chemistry II Lecture]."

However, other students were challenged by this component of the assignment, and it was evident that these students were not yet familiar enough with the reaction to recognize it in the context of the laboratory course. For example, one student wrote:

"I like that this assignment was a little bit more unique and not just a summary of the experiments that we have done. However, some parts of it were confusing since we haven't directly addressed them in class."

This lack of familiarity was also described in the interviews (e.g., the excerpt from Madeline's interview in 5.11.3 Appendix 3, Table 5.4). That not all students recognized the content within this assignment that was formally introduced in the concurrent lecture course highlights the essential role instructors have for designing instruction and assignments that explicitly connect to students' prior knowledge. Although the goal of this assignment was to relate to students' prior knowledge by asking them to describe a familiar mechanism and to build upon that knowledge by designing an analogue and discussing NMR, it was evident that some students did not recognize the acid hydrolysis mechanism in the context of the assignment. This lack of recognition suggests the importance of assignment features to help students recognize these connections.

5.7.2 What components of the writing-to-learn assignments do students perceive as supporting their learning of organic chemistry course content?

Information about the assignment components that support students' meaningful learning was present in the coding for all feedback survey questions. The findings are presented as the themes identified during the analysis: specifically, how the assignment supported students' meaningful learning by (1) *encouraging problem-solving*; (2) *including specific rhetorical components*; (3) *having clear expectations, support, and resources*; and (4) *engaging students in the peer review process*. These themes are summarized in Table 5.2 and described in detail below, with identification of how different assignment components supported students' meaningful learning.

Table 5.2. Themes related to RQ2: What components of writing-to-learn assignments do students perceive as supporting their meaningful learning of organic chemistry course content?

Theme	Exemplar
Encouraging problem-solving	*Thalidomide feedback survey:* "I thought the assignment was quite engaging and required me to learn more about racemization and acid hydrolysis that I had not known previously."
Including rhetorical assignment components	*Acid-base feedback survey:* "I liked the context for this assignment. I thought that rather than writing an essay/short response, writing an email to a 'colleague' helped broaden my writing style and was much more interesting to do."
Providing clear expectations, support, and resources	*Wittig feedback survey:* "I like that the instructions are detailed and well written so that we know what questions to answer when thinking about our response. It really helps narrow down which information to include."
Engaging students in the peer review process	*Acid-base feedback survey:* "While it kind of feels like a hassle to have to review three other writing assignments, I enjoyed it much more than I expected to. Reviewing the other assignments helped me to understand the concepts of the problem more than I previously had, and being able to read the revisions for my assignment made me feel more confident in the work I had done/more certain about the work that still needed to be done."

Encouraging problem-solving. This theme includes instances where students described perceiving that the WTL assignments required them to engage with the problems posed by the assignments. The theme is characterized by students indicating the way the assignments required them to solve problems with thought and creativity. For example, one student responded to the acid–base assignment feedback survey with: "I liked that there was a lot of autonomy to the paper that allowed me to talk about multiple scenarios for the acid base reaction." Other students indicated disliking the challenging nature of the assignments; for example, one student's survey response after the acid–base assignment included: "It was overall very challenging which is good to a point but I think it was too challenging."

Students made similar comments about the prompts encouraging problem-solving within the interviews, as exemplified by the excerpt from Lesley's interview in 5.11.3 Appendix 3, Table 5.4. This theme provides further support for the interconnections between the cognitive and affective domains of learning experiences.[10] Additionally, this theme captures the mixed affect students felt about being challenged by the assignments. Students' varied experiences with valuing how the WTL assignments encouraged problem solving highlights how students' affectual experiences depend on the individual students, as suggested in prior research.[8,18,55] Furthermore, that some students valued tasks that challenged them to solve problems aligns with the literature on the quality of students' academic motivation, particularly in how such tasks can support students' autonomous motivation.[23] Overall, this theme suggests that the problems posed by assignments, and the different levels of difficulty for different students, influences students' affective learning experiences. This finding extends the prior research suggesting the inherent relation between students' attitudes and their motivation towards learning.[2]

Including rhetorical assignment components. The assignments' rhetorical components include the genres in which students were constructing their responses, the audiences to whom students were writing, and the context in which they were providing their responses. The prompts' connections to authentic applications appeared within students' responses in a variety of ways, including indications that the assignments helped them to identify why the organic chemistry content might be useful to fields of interest for potential future careers. This is exemplified by a response to the acid–base assignment feedback survey:

> "I liked the application of my organic chemistry knowledge to pharmaceuticals and biochemistry. I am interested in these fields, and it was great to understand some aspects of such a complex field."

Similar comments about relevance were made about all three assignments, as each dealt with a practical scenario in some way. For example, a representative response from the thalidomide assignment feedback survey indicated recognizing the practicality of the organic chemistry content: "I liked hypothesizing different forms of thalidomide that might prevent the teratogenic effects. It was nice seeing practical implications of our chemistry work."

These sentiments were also present within the interviews, as seen in the excerpt from Jessie's interview provided in 5.11.3 Appendix 3, Table 5.4. Students' recognition of the different ways the content of each assignment was relevant to them illustrates how WTL assignments can appeal to a mixture of the individual, societal, and vocational relevance domains.[19] Furthermore, this finding suggests that the WTL assignments successfully incorporated context within assignments that students generally perceived to also support their conceptual learning. This illustrates how assignments implemented throughout the semester can serve to incorporate authentic contexts in ways similar to context-based curricula.[24,25]

Some students found the other rhetorical components of the WTL assignments, specifically framing their essay for the appropriate audience and writing within a specific context, more challenging. These experiences reflect challenges with balancing the level of detail with which they were expected to write about concepts. For example, one student's feedback survey indicated:

> "I found it challenging to work through an entire process and describe it all in paper step-by-step. I can understand it myself, but writing it down makes it more complicated because I do not realize when something needs an explanation and when it does not."

While students found it a challenge to balance the depth of explanation for particular concepts, this reflects the learning goal for the assignments to engage students with content by constructing explanations. Altogether, these responses provide evidence that the rhetorical assignment components that connect the chemistry content to authentic situations did not necessarily interfere with the primary learning objective of the WTL assignments, but rather served to promote some students' interest and engagement. Stimulating interest is necessary for supporting meaningful learning in that students' interest is inherently tied to their motivation for actively incorporating new knowledge into their existing conceptual frameworks.[14,21,22]

Providing clear expectations, support, and resources. Several students described challenges and positive or negative affective experiences surrounding the expectations, support, and available resources for the assignments, relating to the necessity for successful WTL assignments to provide clear expectations.[38,39] Negative affective responses to the feedback surveys reflected what students found unclear about the assignments, though these responses were balanced by students who indicated, with positive affect, that the assignments were clear. These students discussed clarity in terms of both assignment directions and content-specific prompt components. For the assignment directions, students specifically reported finding the acid–base assignment, but not the other two assignments, unclear with regard to the level of detail they were supposed to include in their writing. For example, one student responded to the acid–base feedback survey saying:

> "However, the prompt and general directions for this writing assignment were unclear, and the expectations were not outlined clearly—it was hard to know what to write about, and what was expected of us."

The interview analysis revealed that this sentiment, for some students, arose from the lack of a clear list of expectations enabling them to know what to include in their response (e.g., the excerpt from Gabriella's interview in 5.11.3 Appendix 3, Table 5.4). The lack of clarity from the students' perspective might also be related to the language in the assignment, how the assignment was introduced to students, or the fact that the acid–base assignment was the first WTL assignment. The distinction between these possibilities were not evident in our analysis but would be worthy of future research. Nevertheless, students' feedback echoes the sentiment that clear expectations for writing are necessary for effective WTL assignments.[38,39] Furthermore, our findings about students' experiences with the assignment expectations is necessary to understand for future implementations of WTL assignments, as elements of the teaching environment surrounding the expectations for students are known to influence students' academic motivation.[21,22]

The feedback survey responses identifying that the Wittig and thalidomide WTL assignments were unclear more closely linked to content rather than the assignment directions. For example, a comment from the Wittig feedback survey indicated that the student was not sure about directions expressly asking them to explain the role of a specific reactant:

> "I think that the wording of the 3rd checklist point in the prompt (The role of PBu_3 in Scheme 2 should be explained) was vague and hard to understand—I had to clarify it with a writing fellow."

Similarly, for the thalidomide assignment, students expressed that the expectations for proposing an analogue were not clear: "It's not clear what is close enough to thalidomide to be an analog." In these cases, students appeared to be aware of the assignment expectations themselves but were unsure about how to respond to the questions posed by the assignments.

Some students also wrote about the assignments being challenging because they required them to utilize resources—including peers, writing fellows, or instructors—to complete the assignments. These reports provide evidence that the challenges students had with the assignments promoted social interactions during the writing process, another component of successful WTL assignments.[38,39] Other prompt features that students wrote about in their feedback survey responses with less frequency include the figures, word limit, expectations for citations, and time required to respond to the assignments. As the range of comments associated with assignment expectations suggests, the various expectations, support, and resources surrounding the assignment are as important for creating meaningful learning experiences as the content and rhetorical components of the assignments. These findings corroborate previous research suggesting that clear writing expectations for students are vital in engaging students in meaningful writing experiences.[38,39] Furthermore, this finding extends research indicating how the teaching environment—particularly clear directions and availability of support—influences the quality of students' motivation and their affective experiences.[23]

Engaging students in the peer review process. Students indicated both positive and negative affective experiences with the peer review process. This theme was most prevalent in students' feedback responses to the first WTL assignment, but the topic came up in interviews for all assignments (e.g., the excerpt from Stephen's interview in 5.11.3 Appendix 3, Table 5.4). Students generally discussed the value of peer review and how it helped them with the assignments. Comments about the peer review process included that both receiving and providing feedback helped them understand the content in the assignment better. As indicated by one student's response to the thalidomide feedback survey:

> "I like the unique approach of having other students comment on your assignment. This helps if you were mistaken in some concept because students can practice identifying errors and the one making the error can correct it."

Participating in the peer review process provided students with the reassurance that they included correct conceptual information in their written responses, thereby engaging students in both the affective and cognitive domains of meaningful learning. This finding also suggests the importance of peer review for providing students with peer models and timely feedback, both of which are suggested to enhance students' academic motivation.[21,22] Students ascribed value to both receiving and providing feedback, corroborating related studies that likewise examined the role of each component of the peer review process.[56–58] Furthermore, that students reported being able to successfully engage with concepts during peer review extends the results from prior studies that had similar findings through analyzing students' writing and peer review comments.[30,32,33,36,37]

While many of the students who mentioned peer review did indicate finding the peer review process helpful, a small number reported finding the process unhelpful or challenging. For example, one student expressed not receiving any constructive criticism from the peer review for the acid–base assignment: "… I also wasn't very happy with the peer review; some of my reviewers were rude and unhelpful, and no constructive criticism was given."

Generally, students indicated finding peer review unhelpful when reviewers gave non-constructive feedback or when students received conflicting sets of feedback. Future research should seek to understand ways to support students' abilities to provide constructive feedback and to respond to feedback that is conflicting or not constructive. However, the peer review process generally has been shown to have positive effects for students, even when compared to receiving feedback from content experts.[59–62] Our findings suggest that, although some students may have negative experiences with peer review, many students have positive experiences and the implementation of peer review nevertheless provides the structures for peer interactions that support motivation and encourage reflection and revision.[21,38,39]

5.8 Conclusions

This study presents a qualitative analysis of students' feedback survey responses for three WTL assignments that were implemented in an organic chemistry laboratory course to enhance students' meaningful learning experiences. This research provides the first step in understanding the utility of WTL assignments to facilitate students' meaningful learning experiences within organic chemistry laboratory courses by (1) identifying varied ways that WTL assignments can connect new concepts to students' prior knowledge and (2) identifying how components of WTL assignments and implementation can support students' learning experiences. This research was conducted through the lens of meaningful learning theories, with attention to the literature on developing relevancy through authentic tasks and considering students' academic motivation. This is necessary because research in the chemistry education literature has identified that students' approach to learning is closely related to their perceptions of the course and its relevance.[2,10] While faculty often seek to emphasize the relevance of course material through laboratory components of chemistry courses, there appears to be less focus on affective goals for instruction in organic chemistry and other advanced chemistry laboratories.[15] By presenting necessary insight into students' perceptions of a pedagogy that is designed to support students' affective experiences, writing-to-learn, this study extends the research on organic chemistry students' meaningful learning experiences.

Findings from this study indicate that the WTL assignments implemented in the second-semester organic laboratory course successfully provided students with opportunities for meaningful learning. Students perceived the three assignments to connect to their prior knowledge in slightly different ways: (1) by connecting to ideas from previous courses such as general chemistry and first-semester organic chemistry, (2) by connecting to ideas from the second-semester laboratory course itself, and (3) by connecting to ideas from the concurrent second-semester lecture course. In these ways, the assignments appealed to the cognitive domain of meaningful learning by requiring students to draw from knowledge from both previous and concurrent courses. In particular, students perceived the content of the prompts to challenge them to build connections to their existing knowledge of topics ranging from acid–base chemistry to reaction mechanisms.

The assignments also appealed to the affective domain necessary for meaningful learning by encouraging problem-solving; including rhetorical components that emphasized the relevance of the content; having clear expectations, support, and resources; and incorporating peer review. Our findings relating to each of these themes indicate an interplay between the cognitive and affective domains. Additionally, students' affective experiences appeared to be fostered by the rhetorical framing of each assignment within authentic contexts. Students found these contexts to be relevant to their lives and career goals. Importantly, students generally did not find writing within particular rhetorical contexts to be difficult beyond the challenge of communicating content knowledge, thereby allowing students to grapple with content rather than context. Students' affective experiences were also found to be influenced by the clarity of expectations. Specifically, our results suggest that clear expectations can serve to improve students' affective engagement with the assignments and, when unclear, can hinder engagement. Lastly, the incorporation of peer review generally enhanced students' experiences with the assignments by providing reassurance or guidance about their conceptual understanding. These findings illustrate how the specific components of WTL assignments can influence students' affective experiences through creating experiences that emphasize relevancy while supporting academic motivation.

5.9 Limitations

There are several limitations associated with this study. First, this study was completed at a research-intensive institution in the United States, and the findings may not be transferrable to populations of students at other institutions. Additionally, students were not provided any course credit or incentive for responding to the feedback surveys. The lack of incentive could be responsible for the higher survey response rate for the first assignment in relation to the lower response rates for the last two assignments. It could also be a source of self-selection bias, as students may have been more likely to respond if they had strong opinions. The students participating in interviews were also selected on a voluntary basis, which could also have contributed to self-selection bias. Hence, the results from the feedback survey and interview analysis may not be representative of the entire organic chemistry course population at the institution. Furthermore, the feedback survey questions were broad, open-ended, and not directly aligned with the meaningful learning theories, meaning that students may not have provided a complete depiction of their meaningful learning experiences in their responses. Another limitation is that quantitative measures of meaningful learning were not administered to students either before or after completing the WTL assignments. Furthermore, no comparison groups were included. Hence, the results from this study cannot indicate whether students' expectations for meaningful learning in the organic chemistry laboratory course were influenced or fulfilled by completing the WTL assignments. For these reasons, the results of this study are limited in scope to qualitatively identifying aspects of WTL assignments that influence organic chemistry students' meaningful learning experiences rather than measuring students' meaningful learning experiences.

5.10 Implications

5.10.1 Implications for research

This study demonstrates utilizing the theoretical lens of meaningful learning to qualitatively investigate students' experiences with WTL activities. Chemistry education research has largely focused on quantitative studies of meaningful learning,[16] and it is valuable to employ qualitative methodologies to better understand students' experiences. Future research could similarly employ qualitative methodologies to make sense of students' meaningful learning experiences within other curricula or in response to other pedagogical interventions. Because the present study primarily relied on open-ended survey responses with supplementary interview data for triangulation, future qualitative research on WTL pedagogies could additionally employ alternative methodologies, such as the word lists meant to elicit affective responses used by Galloway et al.,[10] to ascertain a more comprehensive view of students' experiences with WTL assignments. Studies should also be conducted that employ mixed methods to understand students' experiences qualitatively while quantitatively measuring students' meaningful learning in response to pedagogical interventions, using instruments such as the MLLI. Future research could also more directly examine the role that instructors, including the course instructor, teaching assistants, and undergraduate writing fellows, have in contributing to students' meaningful learning experiences with WTL assignments. For instance, prior research by Flaherty et al. suggests that when graduate teaching assistants were trained in a meaningful learning pedagogy, the number and quality of interactions between the teaching assistants and students increased.[63] Similar research should be conducted for pedagogical interventions, including WTL, that are specifically designed to support students' meaningful learning experiences.

5.10.2 Implications for practice

This study suggests that instructors should set clear learning goals when designing writing assignments and should be intentional when considering how new concepts targeted by writing assignments will connect to students' existing knowledge. Our findings suggest that students recognize the elements of writing assignments that are necessary for meaningful learning experiences, particularly how assignments explicitly connect to their knowledge from prior courses, knowledge from the course in which the assignment is given, or knowledge from courses taken concurrently. These findings also imply the recommendation that instructors incorporate rhetorical framing within an authentic context and include structures such as peer review and revision when designing and implementing writing assignments. Students' engagement with authentic contexts and peer review can support both the affective and cognitive domains of the learning experience. Furthermore, this research suggests that clear expectations within the writing assignment are recognized by students and can likely influence their meaningful learning experiences. The rhetorical prompt components and clarity of expectations can influence students' engagement with the assignments, if and how they build connections between concepts, and if students find the assignments relevant or motivating. Details of assignments, such as the presentation of figures, the requirement and format of citations, and the directions provided within assignments, all influence students' experiences with the assignment. By carefully attending to each detail of assignments, instructors can influence students' meaningful learning experiences and thereby encourage students to engage in the process of incorporating new knowledge and ideas into their existing knowledge framework.

5.11 Appendices

5.11.1 Appendix 1. Full text of the three writing-to-learn assignments

Acid–base WTL assignment

Levothyroxine, which is used to treat hypothyroidism, is less effective when taken in combination with calcium carbonate. In contrast, calcium citrate, which is also an over the counter calcium supplement, causes little or no interference with the absorption of levothyroxine by the body (Figure 5.1).

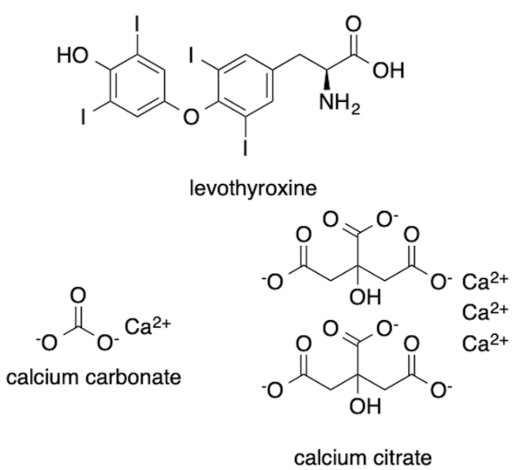


Figure 5.1. Acid–base assignment figure.

For this assignment, you'll take the role of a medicinal chemist writing an email to a collaborator, who is a physician-researcher planning a clinical trial. The goal of the trial is to investigate the co-administration of levothyroxine with calcium supplements. The physician, who took organic chemistry several years ago, has requested your help to clarify the interactions between molecules like levothyroxine and other drugs.

Specifically, your collaborator needs to understand how neutral levothyroxine is deprotonated to its sodium salt form, which is more absorbable in the body. They also need to know how Ca^{2+} as a Lewis acid may interact with levothyroxine to prevent its absorption. Finally, the physician should understand why Ca^{2+} in calcium carbonate may interact with levothyroxine and prevent its absorption whereas Ca^{2+} in calcium citrate will not.

Items to keep in mind:

- This should be an email of between 500–700 words.
- Be sure to explain the relative acidity of each site on levothyroxine and which site would be deprotonated first to make the sodium salt—the form that is given to patients.
- Consider the pH of the stomach acid (1.5 to 3.5) when predicting the predominant Levothyroxine species.
- Remember to appropriately format your email as a letter with a salutation, closing, and proper paragraphing.
- Since you are imagining that you are writing to a colleague, carefully edit and proofread your essay to maintain credibility and consider, as a medicinal chemist, how you would write to an audience in a different field.

Wittig WTL assignment

Recently a research article was published reporting the first base-free catalytic Wittig Reaction. The finding has important implications for industry because the Wittig reaction can be used to make important chemical products on an industrial scale. For example, BASF (the world's largest chemical company) began using the Wittig reaction as a key step in the production of ß-carotene (vitamin A) in the early 1960s. The general Wittig reaction is shown below in Scheme 1, followed by a successful example of the new Wittig reaction from this publication in Scheme 2 (Figure 5.2).

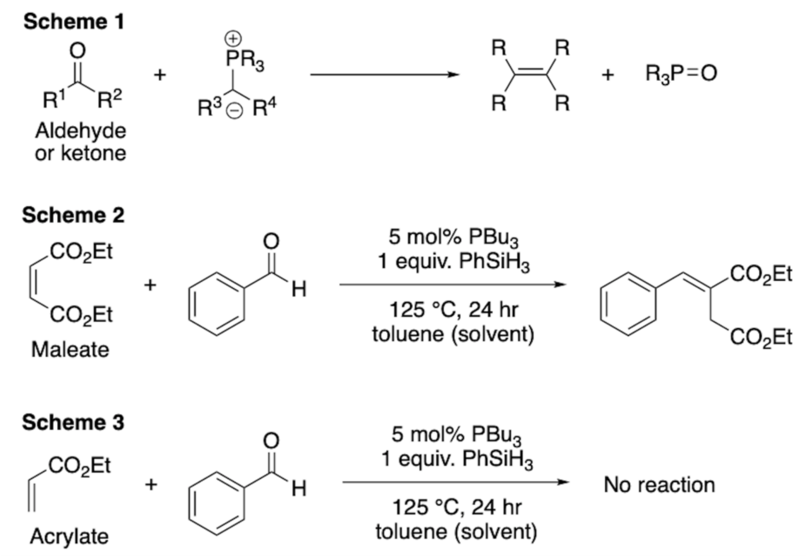


Figure 5.2. Wittig assignment reaction schemes.

You are a science reporter who regularly contributes short pieces that highlight important chemistry discoveries to *Chemical and Engineering News* (C&EN), which is the premier magazine of the American Chemical Society. This research article captured the attention of your editor at C&EN, and she has assigned you to write a highlight about the chemistry described in the study for the upcoming issue of the magazine.

The challenge of highlighting research for C&EN is the wide range of its readership, which includes ~160,000 members who are either professional chemists, chemistry students, or persons in areas that may relate to chemistry and work in academia, industry, non-profits, or policy. Their specialties are wide-ranging and include fields like biochemistry, chemical engineering, inorganic chemistry, medicinal chemistry, or even physics (to name a few). This means that each reader will have some general proficiency in chemistry, but that you should not assume a depth of understanding in organic chemistry. Be sure to translate any organic chemistry jargon or terms for the reader. Use a style and language that is accessible to the broad readership of the magazine. Also keep in mind that it is a news magazine so you should have a catchy title and feel free to take some creative license in your writing to make it more engaging.

Your article should be approximately between 350–750 words in length. In writing your article you must address the following points:

1. Explain the key mechanistic steps that lead to this transformation. Focus specifically on the formation of the ylide in Scheme 2. Note that the ylide that is ultimately formed is not shown in this scheme.
2. Explain why no base is needed in Scheme 2 and contrast with the general Wittig reaction in Scheme 1.
3. Explain the role of PBu_3 in the reaction and why it can react with a functionalized alkene (maleate) instead of an alkyl halide (the standard reaction pathway shown in class). You do not need to discuss the role of $PhSiH_3$ in your draft; the $PhSiH_3$ acts as a reducing agent to regenerate PBu_3.
4. Stress the key aspects of the reaction that make it attractive for industrial scale reactions.
5. Offer an explanation as to why the product is formed in *E/Z* ratio 96/4.
6. Address the limitation (Scheme 3) that the reaction does not work with acrylates (offer an explanation as to why this is so).

Thalidomide WTL assignment

Thalidomide was widely used after World War II as a sedative and later as a treatment for morning sickness. Unfortunately, it was only after widespread use that it was discovered that thalidomide causes very serious side effects – in particular, birth defects such as phocomelia (limb malformation). The drug was banned in 1962 and these events resulted in important changes to the way the FDA approves drugs.

Despite the inherent dangers, thalidomide is now used for treatment of serious diseases, such as cancer and leprosy, when the benefit of treatment outweighs the inherent risks. It is now understood that thalidomide exists as two enantiomers; one is a teratogen and the other has therapeutic properties. Rapid racemization occurs at body pH and both enantiomers are formed at roughly an equal mixture in the blood, which means that even if only the useful isomer is used, both will form once introduced in the body. Furthermore, both enantiomers are subject to acid hydrolysis in the body and produce hydrolysis products that may or may not be teratogens depending on their structure. The structure of thalidomide and two thalidomide hydrolysis products are shown below in Figure 5.3.

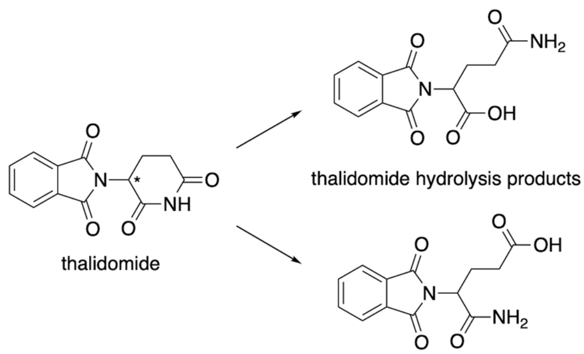


Figure 5.3. Thalidomide assignment Fig. 1.

You are an organic chemist collaborating with a team of other researchers from the University of Michigan with the goal of testing thalidomide analogs for cancer treatment. An analog is a compound that is very similar to the pharmaceutical target that has small structural differences. For example, *m*-cresol (shown in Figure 5.4 below) is an analog of phenol.

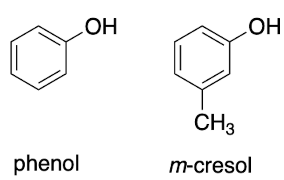


Figure 5.4. Thalidomide assignment Fig. 2.

Your goal will be to design a structural difference that will make the thalidomide analog less reactive toward hydrolysis than thalidomide. Your analogs will be tested for the inhibition of a pro-inflammatory protein mediator, which in elevated levels may be responsible for symptoms associated with the early stages of HIV.

Although thalidomide is warranted for treatment of some diseases, it would be preferable to identify an analog that has similar therapeutic qualities without the potentially devastating side effects. It is known that thalidomide is easily hydrolyzed, and it has been proposed that one of the biologically active species may be one of the two possible hydrolysis products shown above in Figure 5.3. Thus it is important to propose analogs that are not readily hydrolyzed.

Your research team is drafting a grant proposal for the National Institute of Health. You must contribute between a 350–750 word description explaining the structure and reactivity of thalidomide toward hydrolysis and the structural differences in proposed analogs that will make them inert to hydrolysis. Set the tone of your piece by placing your description in the context of the larger goal of developing a safer drug for the treatment of cancer patients. The committee who will review the proposal is likely to be made up of scientists from disciplines including biology, chemistry and medicine. While they are experts in their own field, they may not be knowledgeable about organic chemistry, racemization, hydrolysis, or NMR spectroscopy. You should consider carefully which organic chemistry terms you use and when you define or explain them. Remember, your collaborators are relying on you to clearly communicate your plan so that they can write a competitive proposal for funding from the NIH.

When writing, you should consider the following:

1. Explain the mechanism for acid hydrolysis of thalidomide to form the two hydrolysis products in Figure 5.3.
2. Design one compound (thalidomide analog) that should be a pro-inflammatory protein mediator inhibitor. Explain. Keep in mind that any changes to the structure of a molecule can result in vastly different activity in the body.
3. Explain why it is important that thalidomide analogs do not have acidic protons at their stereocenters.
4. Describe how you would monitor hydrolysis of thalidomide by NMR.
5. Be sure to cite any outside sources that you used while writing your paper. Images that you did not draw yourself must have the original source cited. Sources should be cited using the APA/ACS format.

Note: You can choose to include drawings of either the mechanism or of your proposed analog. However, given your audience, your written explanation should be sufficient such that your proposed analog can be understood without the drawing.

5.11.2 Appendix 2. Complete coding scheme

Table 5.3. Complete coding scheme

Code (Survey Question)	**Definition**	**Exemplar**	**Frequency, Acid-base ($N = 333$)**	**Frequency, Wittig ($N = 149$)**	**Frequency, Thalidomide ($N = 147$)**
Theme: Building connections between content					

New concept (Q2)	The student indicated being challenged by new concept introduced by the assignment.	"The hardest part was figuring out why the reaction doesn't need a base and figuring out that there must be a proton rearrangement that goes on."	228	110	120
Conceptual understanding (Q1, +)[a]	The student indicated feeling that the assignments helped them develop conceptual understanding.	"The assignment challenges me to understand the organic chemistry concepts and articulate them properly/clearly."	16	19	10
Relevant to class (Q1, +)[a]	The student indicated that the assignment material related to the course content.	"I liked this assignment because it was related to what we were doing in lab and seemed more relevant than [the acid-base assignment]."	70	37	20
Not relevant to class (Q1, –)[a]	The student indicated feeling underprepared to respond to the assignment based on the course content.	"The presentation was pretty clear but we were not equipped with insight from lecture to know how to answer the questions."	22	11	13
Prior knowledge (Q1, +)[a]	The student indicated drawing from previous knowledge and applying it to new situations.	"It brought a reaction which was talked about in [Organic Chemistry I Lecture] (hydrolysis) and connected it to [Organic Chemistry II Lecture and Laboratory] material."	30	3	1
Theme: Encouraging problem solving					
Problem-solving (Q1, +)[a]	The student indicated feeling satisfied by the problem-solving required to complete the assignment.	"When I was able to make sense if an unknown reaction I was very satisfied."	62	38	45
Thought-provoking (Q1, +)[a]	The student indicated that the assignment was thought-provoking.	"I liked that it required thought outside the actual lab pages results."	45	33	23
Creative (Q1, +)[a]	The student indicated feeling satisfied by the assignments supporting their creativity.	"The assignment allowed us to be creative while still learning chemistry along the way."	16	13	15

Challenging (Q1, +)[a]	The student indicated feeling challenged by the assignment in a productive way.	"I liked the challenge of considering the pK_a of several molecules in order to explain how they interact."	1	5	1
Challenging (Q1, –)[a]	The student indicated feeling challenged by the assignment in an unproductive way.	"At first I thought it was hard to approach, and without office hours I would be very lost on how to answer some questions."	16	5	12
Easy (Q1, +)[a]	The student indicated finding components of the assignment easy to complete.	"The mechanism was pretty obvious, as were the changes that would be observed in the H NMR"	1	3	3
Theme: Rhetorical assignment components					
Relevance (Q1, +)[a]	The student indicated feeling that the assignment related content to a practical, authentic situation.	"I also liked how the assignment asked us to relate this new reaction to real-life industrial applications."	31	40	31
Audience (Q2)	The student indicated being challenged by writing to the designated audience.	"At first, I also struggled with organizing my email for a person not in organic chemistry to better understand."	22	7	2
Context (Q2)	The student indicated being challenged by the level of detail with which they were expected to write about concepts.	"I found it challenging to work through an entire process and describe it all in paper step-by-step. I can understand it myself, but writing it down makes it more complicated because I do not realize when something needs an explanation and when it does not."	72	31	29
Theme: Providing clear expectations, support, and resources.					
Clear (Q1, +)[a]	The student indicated that the assignment (or part of the assignment) was clear to them.	"The goals of the writing assignment and the questions asked about it were clear."	66	25	17

Unclear (Q1, –)[a]	The student indicated that the assignment (or part of the assignment) was unclear to them.	"The only thing that was a little unclear at first was considering why one product was major over the other one."	92	25	45
Prompt directions (Q2)	The student indicated being challenged by the directions in the prompt.	"It was very challenging to figure out exactly what each question was asking and to answer it in the way that they wanted."	35	16	14
Outside help (Q2)	The student indicated needing to utilize resources to complete the assignment, including internet searches or seeking help from graduate student instructors, writing fellows, or peers.	"I thought that it was very difficult to come up with the ideas by myself or through research, only after talking it through with friends did I understand what was going on."	31	17	12
Figure (Q1, +)[a]	The student indicated finding the figures helpful.	"I liked that the schemes were included in the prompt; it made it a lot easier to understand the differences between each."	1	7	3
Figure (Q1, –)[a]	The student indicated problems with reading the figures.	"I did not like how the diagrams given were blurry and difficult to read."	1	3	0
Word limit (Q1, –)[a]	The student indicated feeling constrained and/or challenged with the word limit.	"Not much about the situation/prompt was unclear, but I was unsure about the word count, since it was 'approximately 350-750' words. Does that mean that we cannot write more than that? Or is it acceptable if we write close to that range?"	3	1	1
Citing (Q1, –)[a]	The student indicated problems with providing citations.	"For the first draft, it was unclear what images needed to be cited."	2	3	1
Time (Q1, –)[a]	The student indicated feeling the assignment took more time than was needed.	"I did feel that the assignment was a little bit of a waste of time. I felt that my time was needed in other ways for my success in this class."	6	5	5

Theme: Engaging students in the peer review process					
Peer-review (Q1, +)[a]	The student indicated feeling that the peer-review process was helpful.	"I thought that the peer revision phase of this assignment was really helpful in making sure I got my ideas across."	81	4	8
Peer-review (Q1, –)[a]	The student indicated feeling that the peer-review process was unhelpful.	"I also wasn't very happy with the peer review; some of my reviewers were rude and unhelpful, and no constructive criticism was given."	8	1	0
Peer-review (Q2)	The student indicated challenges with the peer-review process.	"...the peer reviews that I read and received all had different understandings of the prompt. I also found it challenging to translate my peer review feedback into my revisions because I received three very different sets of feedback… getting very polar feedback created a challenge for me."	16	2	2
[a] Codes with a "+" applied to responses with positive affect, while codes with a "–" applied to responses with negative affect.					

5.11.3 Appendix 3. Interview excerpts

Table 5.4. Interview responses relating to themes that emerged from the feedback survey analysis

Theme	Exemplar
Themes related to RQ1: How do organic chemistry students perceive building connections between new concepts and their existing knowledge when responding to the writing-to-learn assignments?	
Building on prior knowledge	*Gabriella, acid-base assignment interview:* "…it made a pretty good review of [Organic Chemistry I Lecture] ideas going into [Organic Chemistry II Lecture, because this was assigned during just the very beginning." *Matthew, acid-base assignment interview:* "…it felt like out of left field. We weren't doing this in lab. We hadn't done this in [Organic Chemistry II] lecture. It was just really random. I'll write about it, because I kind of know what I'm talking about, but at the same time…"
Building on course concepts	*Jameson, Wittig assignment interview:* "This is something definitely I should have been able to figure out but it's not like, you know, like a sort of stale question. Like, okay, what is the Wittig reaction or blah, blah, blah. Because you can just copy that out of your notes. So these were the two questions. . . that actually we had to sort of think and say okay, well, we have to use our logic here and see, okay, why wouldn't there need to be a base. Or, why won't this reaction work in this situation?"
Building on concepts from a concurrent course	*Madeline, thalidomide assignment interview:* "...for this assignment… I couldn't tie it to one lecture that we took in class or something we did in lab. I don't know why. I know we did it, but it was ... I think the other assignments were pretty straightforward, like they have the same title as the lab that we're doing in class."

Themes related to RQ2: What components of writing-to-learn assignments do students' perceive as supporting their meaningful learning of organic chemistry course content?	
Encouraging problem-solving	*Lesley, thalidomide assignment interview:* "I thought it was kind of interesting, it was almost like a puzzle, figuring out how can you change this to halt the process. That was a new way to look at, to examine a reaction. How you can prevent the reaction from going past a certain point. What can you do to stop that? It made me look at reactions in a slightly different way too."
Including rhetorical assignment components	*Jessie, Wittig assignment interview:* "I never thought about it because we always think about organic chemistry in class and on paper. I never think of it as real people are using it to make real things, which I probably should… It was kind of cool seeing chemistry in this light, where more as like it's a product that's being bought and sold, and they needed to save money."
Providing clear expectations, support, and resources	*Gabriella, acid-base assignment interview:* "Maybe clarifying expectations with the calcium carbonate/citrate differentiations. That was something that I wasn't sure if I was supposed to do extensive research, and I decided to just kind of get some sort of an idea by looking at other publications... So specifying that it's something that we should be able to infer based off the structural information would probably have been helpful."
Engaging students in the peer review process	*Stephen, Wittig assignment interview:* "I think it's nice for an outside reader to say when something's confusing and... and a lead on why so that I can work on making things easier to follow, because I do feel like when you're talking about electrons moving around in a reaction, it's easy to make something hard to follow probably. Those kinds of feedback were helpful."

5.12 Acknowledgements

The authors would like to thank the University of Michigan Third Century Initiative and the Keck Foundation for funding. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260. The authors would also like to thank Atia Sattar, Ashley Karlin, and Barry C. Thompson for their input on constructing the three writing-to-learn assignments, and the members of the Shultz group for discussions related to the analysis and results presented in this manuscript.

5.13 References

(1) Anderson, T. L.; Bodner, G. M. What Can We Do about "Parker"? A Case Study of a Good Student Who Didn't "get" Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–101. https://doi.org/10.1039/b806223b.

(2) Grove, N. P.; Bretz, S. L. A Continuum of Learning: From Rote Memorization to Meaningful Learning in Organic Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 201–208. https://doi.org/10.1039/c1rp90069b.

(3) Graulich, N. The Tip of the Iceberg in Organic Chemistry Classes: How Do Students Deal with the Invisible? *Chem. Educ. Res. Pract.* **2015**, *16* (1), 9–21. https://doi.org/10.1039/c4rp00165f.

(4) Karty, J. M.; Gooch, G.; Bowman, B. G. Teaching a Modified Hendrickson, Cram, and

Hammond Curriculum in Organic Chemistry: Curriculum Redesign to Turn around Student Performance. *J. Chem. Educ.* **2007**, *84* (7), 1209–1216. https://doi.org/10.1021/ed084p1209.

(5) Grove, N. P.; Hershberger, J. W.; Bretz, S. L. Impact of a Spiral Organic Curriculum on Student Attrition and Learning. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 157–162. https://doi.org/10.1039/b806232n.

(6) Hein, S. M. Positive Impacts Using POGIL in Organic Chemistry. *J. Chem. Educ.* **2012**, *89* (7), 860–864. https://doi.org/10.1021/ed100217v.

(7) Galloway, K. R.; Bretz, S. L. Development of an Assessment Tool to Measure Students' Meaningful Learning in the Undergraduate Chemistry Laboratory. *J. Chem. Educ.* **2015**, *92* (7), 1149–1158. https://doi.org/10.1021/ed500881y.

(8) Galloway, K. R.; Bretz, S. L. Measuring Meaningful Learning in the Undergraduate Chemistry Laboratory: A National, Cross-Sectional Study. *J. Chem. Educ.* **2015**, *92* (12), 2006–2018. https://doi.org/10.1021/acs.jchemed.5b00538.

(9) Galloway, K. R.; Bretz, S. L. Measuring Meaningful Learning in the Undergraduate General Chemistry and Organic Chemistry Laboratories: A Longitudinal Study. *J. Chem. Educ.* **2015**, *92* (12), 2019–2030. https://doi.org/10.1021/acs.jchemed.5b00754.

(10) Galloway, K. R.; Malakpa, Z.; Bretz, S. L. Investigating Affective Experiences in the Undergraduate Chemistry Laboratory: Students' Perceptions of Control and Responsibility. *J. Chem. Educ.* **2016**, *93* (2), 227–238. https://doi.org/10.1021/acs.jchemed.5b00737.

(11) Novak, J. D. Human Constructivism: A Unification of Psychological and Epistemological Phenomena in Meaning Making. *Int. J. Pers. Constr. Psychol.* **1993**, *6* (2), 167–193. https://doi.org/10.1080/08936039308404338.

(12) Bretz, S. L. Novak's Theory of Education: Human Constructivism and Meaningful Learning. *J. Chem. Educ.* **2001**, *78* (8), 1107. https://doi.org/10.1021/ed078p1107.6.

(13) Novak, J. D. Meaningful Learning: The Essential Factor for Conceptual Change in Limited or Inappropriate Propositional Hierarchies Leading to Empowerment of Learners. *Sci. Educ.* **2002**, *86* (4), 548–571. https://doi.org/10.1002/sce.10032.

(14) Ausubel, D. P. *The Psychology of Meaningful Verbal Learning: An Introduction to School Learning*; Grune & Stratton: New York, 1963.

(15) Bretz, S. L.; Fay, M.; Bruck, L. B.; Towns, M. H. What Faculty Interviews Reveal about Meaningful Learning in the Undergraduate Chemistry Laboratory. *J. Chem. Educ.* **2013**, *90* (3), 281–288. https://doi.org/10.1021/ed300384r.

(16) Flaherty, A. A. A Review of Affective Chemistry Education Research and Its Implications

for Future Research. *Chem. Educ. Res. Pract.* **2020**, No. 2015. https://doi.org/10.1039/c9rp00200f.

(17) Enneking, K. M.; Breitenstein, G. R.; Coleman, A. F.; Reeves, J. H.; Wang, Y.; Grove, N. P. The Evaluation of a Hybrid, General Chemistry Laboratory Curriculum: Impact on Students' Cognitive, Affective, and Psychomotor Learning. *J. Chem. Educ.* **2019**, *96* (6), 1058–1067. https://doi.org/10.1021/acs.jchemed.8b00637.

(18) Hensen, C.; Glinowiecka-Cox, G.; Barbera, J. Assessing Differences between Three Virtual General Chemistry Experiments and Similar Hands-On Experiments. *J. Chem. Educ.* **2020**, *97* (3), 616–625. https://doi.org/10.1021/acs.jchemed.9b00748.

(19) Stuckey, M.; Hofstein, A.; Mamlok-Naaman, R.; Eilks, I. The Meaning of "relevance" in Science Education and Its Implications for the Science Curriculum. *Stud. Sci. Educ.* **2013**, *49* (1), 1–34. https://doi.org/10.1080/03057267.2013.802463.

(20) Stuckey, M.; Eilks, I. Increasing Student Motivation and the Perception of Chemistry's Relevance in the Classroom by Learning about Tattooing from a Chemical and Societal View. *Chem. Educ. Res. Pract.* **2014**, *15* (2), 156–167. https://doi.org/10.1039/c3rp00146f.

(21) Schunk, D. H. Self-Efficacy and Academic Motivation. *Educ. Psychol.* **1991**, *26* (3–4), 207–231. https://doi.org/10.1080/00461520.1991.9653133.

(22) Vansteenkiste, M.; Lens, W.; Deci, E. L. Intrinsic versus Extrinsic Goal Contents in Self-Determination Theory: Another Look at the Quality of Academic Motivation. *Educ. Psychol.* **2006**, *41* (1), 19–31. https://doi.org/10.1207/s15326985ep4101_4.

(23) Vansteenkiste, M.; Sierens, E.; Soenens, B.; Luyckx, K.; Lens, W. Motivational Profiles From a Self-Determination Perspective: The Quality of Motivation Matters. *J. Educ. Psychol.* **2009**, *101* (3), 671–688. https://doi.org/10.1037/a0015083.

(24) Gilbert, J. On the Nature of "Context" in Chemical Education. *Int. J. Sci. Educ.* **2006**, *28* (9), 957–976. https://doi.org/10.1080/09500690600702470.

(25) Pilot, A.; Bulte, A. The Use of "Contexts" as a Challenge for the Chemistry Curriculum: Its Successes and the Need for Further Development and Understanding. *Int. J. Sci. Educ.* **2006**, *28* (9), 1087–1112. https://doi.org/10.1080/09500690600730737.

(26) Rivard, L. O. P. A Review of Writing to Learn in Science: Implications for Practice and Research. *J. Res. Sci. Teach.* **1994**, *31* (9), 969–983. https://doi.org/10.1002/tea.3660310910.

(27) Reynolds, J. A.; Thaiss, C.; Katkin, W.; Thompson, R. J. Writing-to-Learn in Undergraduate Science Education: A Community-Based, Conceptually Driven Approach. *CBE Life Sci. Educ.* **2012**, *11* (1), 17–25. https://doi.org/10.1187/cbe.11-08-0064.

(28) Klein, P. D.; Boscolo, P. Trends in Research on Writing as a Learning Activity. *J. Writ. Res.* **2016**, *7* (3), 311–351. https://doi.org/10.17239/jowr-2016.07.03.01.

(29) Shultz, G. V.; Gere, A. R. Writing-to-Learn the Nature of Science in the Context of the Lewis Dot Structure Model. *J. Chem. Educ.* **2015**, *92* (8), 1325–1329. https://doi.org/10.1021/acs.jchemed.5b00064.

(30) Moon, A.; Zotos, E.; Finkenstaedt-Quinn, S.; Gere, A. R.; Shultz, G. Investigation of the Role of Writing-to-Learn in Promoting Student Understanding of Light-Matter Interactions. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 807–818. https://doi.org/10.1039/c8rp00090e.

(31) Moon, A.; Moeller, R.; Gere, A. R.; Shultz, G. V. Application and Testing of a Framework for Characterizing the Quality of Scientific Reasoning in Chemistry Students' Writing on Ocean Acidification. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 484–494. https://doi.org/10.1039/c9rp00005d.

(32) Finkenstaedt-Quinn, S. A.; Snyder-White, E. P.; Connor, M. C.; Gere, A. R.; Shultz, G. V. Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *J. Chem. Educ.* **2019**, *96* (2), 227–237. https://doi.org/10.1021/acs.jchemed.8b00711.

(33) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Kasner, G.; Wilhelm, C. A.; Moon, A.; Gere, A. R.; Shultz, G. V. Capturing Student Conceptions of Thermodynamics and Kinetics Using Writing. *Chem. Educ. Res. Pract.* **2020**, *21*, 922–939. https://doi.org/10.1039/C9RP00292H.

(34) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(35) Watts, F. M.; Schmidt-McCormack, J.; Wilhelm, C.; Karlin, A.; Sattar, A.; Thompson, B.; Gere, A. R.; Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21*, 1148–1172. https://doi.org/10.1039/C9RP00185A.

(36) Halim, A. S.; Finkenstaedt-Quinn, S. A.; Olsen, L. J.; Gere, A. R.; Shultz, G. V. Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review. *CBE Life Sci. Educ.* **2018**, *17* (2), 1–12. https://doi.org/10.1187/cbe.17-10-0212.

(37) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Chambers, T. G.; Moon, A.; Goldman, R. S.; Gere, A. R.; Shultz, G. V. Investigation of the Influence of a Writing-To-Learn Assignment on Student Understanding of Polymer Properties. *J. Chem. Educ.* **2017**, *94* (11), 1610–1617. https://doi.org/10.1021/acs.jchemed.7b00363.

(38) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Res. Teach. English* **2015**, *50* (2), 199–235.

(39) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and Conceptual Learning in Science: An Analysis of Assignments. *Writ. Commun.* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(40) Eodice, M.; Geller, A. E.; Lerner, N. *The Meaningful Writing Project: Learning, Teaching, and Writing in Higher Education*; Utah State University Press: Boulder, Colorado, 2016.

(41) Bulte, A.; Westbroek, H.; de Jong, O.; Pilot, A. A Research Approach to Designing Chemistry Education Using Authentic Practices as Contexts. *Int. J. Sci. Educ.* **2006**, *28* (9), 1063–1086. https://doi.org/10.1080/09500690600702520.

(42) Miles, M. B.; Huberman, A. M.; Saldana, J. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed.; Sage: Los Angeles, CA, 2014.

(43) Herrington, D. G.; Daubenmire, P. L. Using Interviews in CER Projects: Options, Considerations, and Limitations. *ACS Symp. Ser.* **2014**, *1166*, 31–59. https://doi.org/10.1021/bk-2014-1166.ch003.

(44) Kirilenko, A. P.; Stepchenkova, S. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLoS One* **2016**, *11* (3), 1–14. https://doi.org/10.1371/journal.pone.0149787.

(45) McHugh, M. L. Lessons in Biostatistics Interrater Reliability : The Kappa Statistic. *Biochem. Medica* **2012**, *22* (3), 276–282.

(46) Cartrette, D. P.; Mayo, P. M. Students' Understanding of Acids/Bases in Organic Chemistry Contexts. *Chem. Educ. Res. Pract.* **2011**, *12* (1), 29–39. https://doi.org/10.1039/c1rp90005f.

(47) McClary, L.; Talanquer, V. College Chemistry Students' Mental Models of Acids and Acid Strength. *J. Res. Sci. Teach.* **2011**, *48* (4), 396–413. https://doi.org/10.1002/tea.20407.

(48) Stoyanovich, C.; Gandhi, A.; Flynn, A. B. Acid-Base Learning Outcomes for Students in an Introductory Organic Chemistry Course. *J. Chem. Educ.* **2015**, *92* (2), 220–229. https://doi.org/10.1021/ed5003338.

(49) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid-Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(50) Flynn, A. B.; Amellal, D. G. Chemical Information Literacy: PKa Values-Where Do Students Go Wrong? *J. Chem. Educ.* **2016**, *93* (1), 39–45. https://doi.org/10.1021/acs.jchemed.5b00420.

(51) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.; Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions via App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 878–892. https://doi.org/10.1039/c9rp00260j.

(52) 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.

(53) Graulich, N.; Schween, M. Concept-Oriented Task Design: Making Purposeful Case Comparisons in Organic Chemistry. *J. Chem. Educ.* **2018**, *95* (3), 376–383. https://doi.org/10.1021/acs.jchemed.7b00672.

(54) Connor, M. C.; Finkenstaedt-Quinn, S. A.; Shultz, G. V. Constraints on Organic Chemistry Students' Reasoning during IR and 1H NMR Spectral Interpretation. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 522–541. https://doi.org/10.1039/c9rp00033j.

(55) Galloway, K. R.; Bretz, S. L. Using Cluster Analysis to Characterize Meaningful Learning in a First-Year University Chemistry Laboratory Course. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 879–892. https://doi.org/10.1039/c5rp00077g.

(56) Lundstrom, K.; Baker, W. To Give Is Better than to Receive: The Benefits of Peer Review to the Reviewer's Own Writing. *J. Second Lang. Writ.* **2009**, *18* (1), 30–43. https://doi.org/10.1016/j.jslw.2008.06.002.

(57) Cho, Y. H.; Cho, K. Peer Reviewers Learn from Giving Comments. *Instr. Sci.* **2011**, *39* (5), 629–643. https://doi.org/10.1007/s11251-010-9146-1.

(58) Cho, K.; MacArthur, C. Learning by Reviewing. *J. Educ. Psychol.* **2011**, *103* (1), 73–84. https://doi.org/10.1037/a0021950.

(59) Cho, K.; Schunn, C. D. Scaffolded Writing and Rewriting in the Discipline: A Web-Based Reciprocal Peer Review System. *Comput. Educ.* **2007**, *48* (3), 409–426. https://doi.org/10.1016/j.compedu.2005.02.004.

(60) Patchan, M. M.; Charney, D.; Schunn, C. D. A Validation Study of Students' End Comments: Comparing Comments by Students, a Writing Instructor, and a Content Instructor. *J. Writ. Res.* **2009**, *1* (2), 124–152. https://doi.org/10.17239/jowr-2009.01.02.2.

(61) Zhang, F.; Schunn, C. D.; Baikadi, A. Charting the Routes to Revision: An Interplay of Writing Goals, Peer Comments, and Self-Reflections from Peer Reviews. *Instr. Sci.* **2017**, *45* (5), 679–707. https://doi.org/10.1007/s11251-017-9420-6.

(62) Cho, K.; MacArthur, C. Student Revision with Peer and Expert Reviewing. *Learn. Instr.* **2010**, *20* (4), 328–338. https://doi.org/10.1016/j.learninstruc.2009.08.006.

(63) Flaherty, A.; O'Dwyer, A.; Mannix-McNamara, P.; Leahy, J. J. Evaluating the Impact of the "Teaching as a Chemistry Laboratory Graduate Teaching Assistant" Program on Cognitive and Psychomotor Verbal Interactions in the Laboratory. *J. Chem. Educ.* **2017**, *94* (12), 1831–1843. https://doi.org/10.1021/acs.jchemed.7b00370.

Chapter 6
Investigating Writing-To-Learn To Elicit Organic Chemistry Sudents' Reasoning About Resonance

6.1 Initial remarks

This chapter presents the first of three studies which analyze students' written responses to different writing-to-learn (WTL) assignments to examine how WTL can elicit students' reasoning in organic chemistry. Each of the three chapters involves a different WTL assignment implemented in the second-semester organic chemistry laboratory course. This chapter focuses on a WTL assignment eliciting students' understanding of resonance (a fundamental concept in organic chemistry) and how resonance influences reactivity. The following chapters examine WTL assignments that elicit students' reasoning with reaction mechanisms. The set of three chapters build on Chapters 4 and 5, which establish how WTL can promote engagement and meaningful learning, by focusing specifically on how WTL can also support students' reasoning. Furthermore, the set of three chapters contribute implications to the organic chemistry education research literature regarding how students understand and reason with these topics in organic chemistry.

The WTL assignment in this study was developed to elicit students' understanding of resonance by asking them to write an *explainer* (i.e., a short article that explains information to a broad audience) about resonance. The assignment included directions for students to describe why resonance is important and to discuss a specific example of the influence of resonance on chemical reactivity. The study was grounded in the cognitive process theory of writing, which describes writing as a process in which a writer recursively engages with a writing task and their internal representations of pertinent knowledge to produce text. The cognitive process theory justifies the analytical decision to examine students' final drafts from the WTL assignment, which were qualitatively analyzed with a framework that describes the essential learning outcomes for the resonance concept. The results revealed various analogies, examples, and definitions of resonance students used in their writing. For example, some students used an analogy of taking pictures of an object from different perspectives, where the different pictures are analogous to resonance contributors. This analogy successfully communicates that resonance contributors provide a different perspective on molecular structure but does not account for the fact that resonance contributors themselves do not exist in reality. Hence, the ways students explained resonance provided insight regarding what aspects of the concept they may find more important. Another key finding was that students described the concept operationally (i.e., with a focus on how to draw different resonance contributors) rather than conceptually (i.e., with a focus on what resonance means at a conceptual level). When describing the influence of resonance on reactivity, students included surface-level explanations of how resonance influences reactivity, often equating resonance with stability. These findings regarding how students conceptualize resonance provide implications for instructors by detailing the nuanced and various ideas students may have about the concept, which can guide how instructors present the topic.

This chapter was originally published as a research article in the *Journal of Chemical Education*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author on the manuscript, I contributed to conceptualization, methodology, data analysis, and writing (both original draft preparation and review and editing). I shared primary authorship with P.B. Brandfonbrener, an undergraduate mentee, who contributed to conceptualization, data analysis, and writing (original draft preparation for sections of the manuscript). G.V. Shultz contributed to project supervision, conceptualization, data collection, and writing (review and editing).

Original publication and copyright information

Reprinted with permission from P.B. Brandfonbrener, F.M. Watts, and G.V. Shultz, *J. Chem. Educ*. 2021, **98**, 3431–3441. Copyright 2021 American Chemical Society.

6.2 Abstract

Resonance is a fundamental concept that is necessary for students' successful learning in organic chemistry. However, there is a need to know more about both (1) what students find important when describing resonance and (2) students' conceptual understanding. This research seeks to address this discrepancy by examining second-semester organic chemistry students' responses to a writing-to-learn (WTL) assignment focused on resonance. This work is guided by

the cognitive process theory of writing and studies within undergraduate STEM and chemistry courses that indicate the usefulness of WTL assignments for examining how students convey their understanding of fundamental concepts. We analyzed students' responses to a WTL assignment designed to elicit students' explanations of resonance and its influence on reactivity to a general, nonscientific audience. The goal of our analysis was to identify the features students found important when explaining resonance and to gain insight regarding their conceptual understanding. The analysis was guided by an analytical framework outlining the learning objectives for resonance at the introductory organic chemistry level. We identified various elements in students' writing, including (1) the analogies or examples students used to explain resonance, (2) the basic definitions of the concept students included, and (3) the various ways students described the influence of resonance on reactivity. Our findings indicate that students generally described resonance with an operational rather than conceptual definition, focusing on the process of drawing resonance structures rather than providing conceptual explanations. Furthermore, we identified that students tended to associate resonance with stability, with many students often extending this association to make overgeneralized statements about the influence of resonance on reactivity.

6.3 Introduction

Resonance is a key topic introduced in first-semester organic chemistry courses. Students are expected to have a thorough understanding of resonance and to be able to use it across applications concerning both the structure and reactivity of molecules. However, there is limited research devoted to how students conceptualize resonance and connect resonance to reactivity, though organic chemistry instructors report resonance as a fundamental but difficult concept for which many students misunderstand key components.[1] Because of this, there is a need to know more about what conceptual features comprise students' conceptualization of resonance. As such, this study seeks to describe students' understanding as presented in their responses to a writing-to-learn (WTL) assignment implemented to encourage engagement with the concept. The goal of this study is to use students' responses to the WTL assignment to identify what features of resonance students choose to incorporate into their explanations as a means to gain insight into their conceptual understanding of resonance and its influence on reactivity.

Existing studies on teaching and learning resonance in organic chemistry provide insight into students' understanding. Taber demonstrated how some students used "resonance" and the phrase “alternation between single and double bonds” interchangeably when describing the structure of benzene.[2] Kim et al. similarly found that the Kekulé representation of resonance may contribute to students’ difficulty understanding the concept.[3] Other studies found that being able to draw resonance structures or hybrids is not necessarily associated with students’ conceptual learning as much as recognizing the limitations of molecular representations.[4,5] To improve the focus on teaching and assessing resonance, Carle and Flynn recently described 10 learning outcomes (LOs) essential to the concept of delocalization, along with describing how the LOs are taught, practiced, and assessed across seven introductory organic chemistry textbooks, end-of-chapter problems, and summative assessments.[6] They highlighted that resonance structures are often described in terms of their relative stability even though these structures do not exist on their own, which might be a cause of some students’ challenges with the concept.

While the existing studies on resonance provide insight into the teaching and learning of resonance, they indicate a need for further research examining students’ conceptualization of resonance and how it influences reactivity. It is of particular interest to gain insight into students’ understanding during their second semester of introductory organic chemistry, a population that has remained understudied with respect to how they conceptualize resonance. It is necessary to study this population, as they have learned about resonance as a foundational concept early in the first-semester curriculum and have practiced applying it to other concepts in organic chemistry (e.g., specific reaction mechanisms). As such, the goal of this study is to examine second-semester organic chemistry students’ responses to a WTL assignment to understand the aspects of the concept they find important when developing written explanations about resonance and its influence on reactivity. By identifying what students choose to include in their own written explanations of the concept, this research aims to provide insight into how students describe resonance and its influence on reactivity.

The central goal of WTL is to support students’ conceptual understanding through writing, and the implementation of WTL in STEM courses specifically has been a recent focus of increased attention in the literature.[7–12] A variety of studies demonstrate the effectiveness of WTL in undergraduate chemistry[13–19] and other STEM courses.[20,21] Studies of WTL specifically in organic chemistry demonstrate its ability to support and elicit students’ understandings of acid−base chemistry and reaction mechanisms.[18,19] Furthermore, WTL assignments support students’ meaningful engagement with organic chemistry course content.[22] These studies, among others,[23,24] demonstrate the value of examining students' written responses to WTL assignments to make sense of students' reasoning and conceptual understanding. As such, the WTL pedagogy is suited for supporting students' learning and providing a rich source of data to access a large sample of students' understanding. In this way, a WTL assignment focused on eliciting students' explanations of resonance is useful for determining what aspects of the concept they find important. Students' responses can provide insight into the features of the concept students deem to be conceptually foundational.

6.4 Theoretical framework

This study is informed by cognitivist perspectives of WTL, which are drawn from early reports of writing as an instructional strategy to support students' learning.[25] WTL itself is grounded in cognitive theories of writing, which recognize that the individual, cognitive processes that occur during writing are situated within a broader sociocultural environment.[10,26–29] This study is specifically guided by Hayes' cognitive process theory that incorporates two major components, the task environment and the individual writer.[30] The task environment accounts for the assignment itself and its genre and audience along with the text the writer has produced and the medium in which they are writing. The individual writer interacts within the task environment while engaging with specific cognitive processes, holding motivations toward the writing task, and utilizing their working and long-term memory to complete the task. The cognitive processes include reflecting, producing text, and interpreting text, aligning with the traditional stages of writing (planning, writing, and revising). However, unlike views of writing that consider these processes to occur in linear stages, Hayes' model recognizes that these processes occur throughout all parts of writing. Alongside describing how working and long-term memory are engaged during these cognitive processes in ways that can support learning, Hayes' model also recognizes the importance of motivation and affect.[30] Hayes' cognitive process theory of writing informed the implementation of the WTL assignment central to this study and justified the analysis of students' revised drafts to ascertain how students understand the concept of resonance, including both how resonance is represented and how it relates to structure and function.[30]

6.5 Research questions

This research aims to understand how students conceptualize resonance and its influence on reactivity as presented in their responses to a WTL assignment during their second semester of organic chemistry. This study was guided by the following research questions:

1. How do second-semester organic chemistry students describe the resonance concept in response to a WTL assignment?
2. How do second-semester organic chemistry students explain how resonance influences reactivity?

6.6 Methods

6.6.1 Setting and participants

The study was conducted at a large, Midwestern research university in a second-semester organic chemistry laboratory course. The course is taken by students who are both chemistry majors and nonmajors, with most of the students majoring in other disciplines (e.g., neuroscience, biomolecular sciences, chemical engineering, etc.). The course consisted of weekly lectures led by faculty and postdoctoral instructors who introduced experiments and procedures. Graduate student instructors led the weekly laboratory to assist students in conducting experiments. Students commonly take this laboratory course along with the second-semester organic chemistry lecture course. The coursework involved keeping a laboratory notebook, taking quizzes, completing worksheets, and responding to three writing assignments (one of which is the subject of this study). The writing assignments accounted for 14% of the total grade, with each assignment worth 4.7%. The study participants consisted of the 316 students enrolled in the laboratory course who completed the WTL assignment described below and consented to participate.

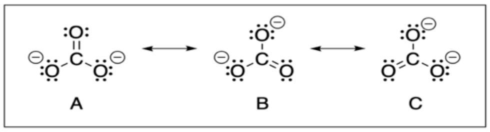
Students were introduced to resonance in their previous semester of organic chemistry and no direct instruction on resonance took place during the lecture of the second-semester course. The first-semester course is cotaught with shared exams, so students entering the second-semester laboratory course were expected to all have a similar introduction to resonance. The first-semester course introduces resonance early, typically within the second and third week of lectures. Resonance is discussed alongside other aspects of drawing and representing organic molecules (e.g., bond-line notation, valency, formal charges, units of unsaturation). The topics specific to resonance that are covered include the limitations of individual Lewis structures and the need for resonance contributors to accurately represent molecular structure. Students are instructed on major/minor resonance contributors and how to determine which resonance structures are more significant. The lectures on resonance also cover identifying when resonance can occur and how to determine all resonance structures for a given molecule. Students are also taught about the necessity for a conjugated system before the course transitions into topics including hybridization and drawing 3D molecular structures. Throughout the course, resonance is emphasized as a foundational concept for considering chemical reactivity. These topics related to resonance are typically covered on the exams for the course and in the practice problems students work to prepare for the exams.

6.6.2 Writing-to-learn assignment design and implementation

The WTL assignment was designed and implemented with attention to the features that support students' conceptual learning.[9,11,12] The assignment development process included content validation with faculty in organic chemistry, English, and education, along with response process validation with undergraduates who had recently taken the second-semester organic chemistry laboratory course. The assignment was the second of three in the semester, intended to draw students' focus to the foundational concept of resonance that was originally introduced at the beginning of the first-semester organic chemistry course. The assignment challenged students to write an *explainer* to educate the general public, assuming a high school education, on the topic of resonance. An *explainer* is a short piece that aims to help the public understand a complex concept using simple language. The instructional objective was to allow students to demonstrate their understanding of resonance and to determine what aspects of this topic are most important, providing an opportunity as researchers to access how students conceptualize resonance. Students were encouraged, but not required, to use an analogy in their response. The assignment also asked students to describe why resonance is important to understand and to discuss a specific example of how resonance influences chemical reactivity. The full prompt can be found in 6.10.1 Appendix 1, and a shortened version of the prompt is presented in Figure 6.1. Each aspect of the assignment sought to engage students in applying their knowledge of resonance during the writing task.[9,11]

Translating Resonance: An Explainer

Consider structures A, B, and C shown below that represent resonance contributors for the carbonate ion. The three structures are not detectable by any experimental method used to examine the structure of carbonate, which indicates that they do not sufficiently describe the carbonate ion.[1] This example demonstrates the limitation of using Lewis structures to depict resonance because none of the structures represent the actual molecule, which is a hybrid of A, B, and C.



In his book, *Thing Explainer: Complicated Stuff in Simple Words*, Randall Munroe, who created the popular web comic series xkcd, challenges his readers to use the most common 1000 words in the English language to explain highly technical concepts. Neil Degrasse Tyson, Carl Sagan, and Bill Nye are well known for these types of translation skills – their ability to conduct rigorous scientific research and talk about why it matters helped make STEM accessible across generations of public audiences. Your objective for this assignment is to write an "explainer" piece that helps a general public understand *resonance*. An "explainer" is a brief radio spot or short journalistic piece, that aims to help a general public access, understand, and potentially apply a complicated or highly technical concept. The explainer often uses an entertaining, simple analogy to make an abstract, complicated idea simpler to understand and visualize. Explainers also give the reader a sense of why a concept is important for them to understand – what can or should we, the general public, do with the information that's been presented to us?

Figure 6.1 A shortened version of the WTL assignment. The full prompt can be found in 6.10.1 Appendix 1. The full version also included examples of explainers and an "Items to keep in mind" section that provided further instructions and guidance for completing the assignment.

Students were given 1 week to submit a first draft of the assignment and were randomly assigned to provide feedback to typically three peers in a double-blind peer review process facilitated by an automated peer review tool. Students were given 3 days to provide feedback by addressing specific content-focused questions, which are provided in 6.10.2 Appendix 2. After receiving feedback, students were required to revise their initial submission before submitting a final draft 1 week later. The content-focused peer review and revision process has been shown to successfully support students in providing feedback on content, rather than grammar or style, in ways that support students' revisions.[16,17,21,31] Throughout the writing process, students had access to writing fellows to receive guidance on the assignment. Writing fellows are a group of undergraduate students, with previous success in the course, who were trained to provide feedback on the WTL assignments. Both the peer review process and the availability of writing fellows served as structures to make the writing task interactive, while the requirement for peer review and revision served to involve students in all components of Hayes' model for the cognitive processes of writing.[9,11,30] These various components surrounding the assignment design and implementation have been shown to support students' meaningful engagement, both cognitively and affectively, with content in an organic chemistry course setting, indicating their role in also supporting the motivational and affective components of Hayes' cognitive process theory of writing.[22,30]

6.6.3 Data collection

The data collected were 316 students' final drafts, from which a randomly selected subset (105 students' final drafts) was analyzed. Before data was collected, approval for the study was granted by the Institutional Review Board and consent was obtained from participating students. The final draft was analyzed in alignment with the cognitive process theory of writing, as it best reflects students' understanding after given time for revision.[30] Specifically, the final draft highlights the aspects of resonance students found most important after completing the peer review process. This draft also reflects the added week of revision time that each student had to decide the most pertinent information to include to fully represent their understanding of resonance in response to the original prompt.

6.6.4 Data analysis

The writing analysis was conducted by developing a coding scheme using a combination of deductive and open coding. The analytical framework for deductive coding was guided by the 10 LOs described in Carle and Flynn's "Essential learning outcomes for delocalization (resonance) concepts: How are they taught, practiced, and assessed in organic chemistry?"[6] The LOs highlight specific areas within the topic of resonance that the authors identified as necessary for students to demonstrate having achieved after taking two semesters of organic chemistry. Our coding scheme used a subset of the LOs from the framework along with additional codes developed through open coding. The full coding scheme and a discussion of its alignment with the analytical framework are provided 6.10.3 Appendix 3 and 6.10.4 Appendix 4. The coding scheme and analytical framework capture the components of resonance students found important for responding to the WTL assignment; applying the coding scheme to students' revised drafts is in alignment with the cognitive process theory of writing for identifying what students found important after engaging in the cognitive processes of writing.[30] Student responses were coded at a document level, and the coding results were used to identify themes across students' writing in alignment with the research questions for this study.[32]

The coding scheme was flexible in its inception and early application, with frequent meetings and discussions among the research team to refine and develop coding definitions using constant comparative analysis.[33] The coding scheme development process took place through multiple stages, whereby two researchers independently coded students' drafts, met to discuss the application of codes, and determined measures of inter-rater reliability (IRR).[34] In the first stages of the analysis process, we followed an inter-rater agreement process in which the application of codes was discussed and a consensus reached after calculating initial reliability measures. The calculated reliability measures included percent agreement, fuzzy κ (a modified version of Cohen's κ that allows for multiple codes to be assigned to each unit of analysis), and Krippendorff's α.[35,36] A total of 35 documents were analyzed during the inter-rater agreement process before reaching saturation and finalizing the coding scheme.[33] With the final coding scheme, two researchers independently coded a reliability subset of 20 documents, achieving 94% agreement with a fuzzy κ value of 0.82, indicating strong agreement.[34] Values of Krippendorff's α for each code are provided in 6.10.3 Appendix 3. The researchers discussed and resolved any discrepancies and then evaluated the original 35 documents to ensure they were analyzed according to the final coding scheme. Afterward, each researcher analyzed an additional randomly selected 25 documents to reach a total of 105 analyzed drafts (33% of all the data). The remaining data was not analyzed as no additional trends were identified in this further analysis, indicating that saturation had been reached.[33]

6.7 Results and discussion

This study aims to understand how second-semester organic chemistry students convey their understanding of the concept of resonance and its influence on reactivity through a WTL assignment. The qualitative analysis of students' written responses first examines how students describe the concept of resonance in general, followed by an analysis of how students use this understanding to explain how resonance influences reactivity. The coding and subsequent analysis are used to report the percentages of students incorporating specific features in their writing in alignment with the identified themes across students' responses.

6.7.1 How do second-semester organic chemistry students describe the resonance concept in response to a WTL assignment?

Analogies students used indicate how they thought about the resonance concept and how resonance is represented. Students were encouraged but not required to include an analogy to explain resonance. For this, 74.3% of students provided an analogy, and the most prevalent analogies give insight into how students may conceptualize resonance. The most common analogy (in 25.6% of responses with analogies) compared resonance contributors to a set of photographs or drawings of the same object from different angles or perspectives (referred to as the picture analogy). Students including this analogy suggested that individual pictures of an object are analogous to resonance structures in that, when these images are examined together or overlaid, they provide a more complete image of the object (analogous to how resonance structures relate to the resonance hybrid). For example, one student wrote,

> "When we take a selfie of ourselves, we have a snapshot of how we look at a particular angle. We can have a hundred pictures of ourselves, each photograph revealing a different angle. However, no single photograph carries all the information about us; we are essentially a hybrid of all the photographs."

In this example, the student suggested that each individual selfie (or resonance contributor) would contribute a different perspective or detail about the actual person (or resonance hybrid) when all the pictures are combined. The student also highlights a strength of this analogy by explicitly stating that no single image (or resonance contributor) can fully describe the object (or molecule). Instead, the resonance hybrid serves as the most accurate representation. Another frequent analogy (in 14.1% of responses with analogies) described resonance as the sharing of electrons between atoms working together as a team to provide stability. One student wrote,

> "Imagine a molecule as a single team. Each atom of the molecule represents a member of the team, and the ball that they pass to one another is a pair of delocalized electrons. On

his own, any one member of the team may feel overwhelmed by his opponents, but with the help of his teammates, he can pass the ball, receive help, and feel the support of his team. Similarly, in a molecule, electrons can be passed between atoms so that no one atom is overwhelmed by too many or not enough electrons."

This analogy, referred to as the teamwork analogy, emphasizes the delocalization of electrons between atoms to stabilize the molecule, a feature of resonance that 98.1% of students directly mentioned in their writing.

In the literature, analogies have been highlighted as a common and effective way for instructors to convey the idea of resonance hybrids.[4] However, little research has examined how students' own analogies may indicate their conceptual understanding. We can explore the differences between the analogies presented in the literature and the analogies students generated to understand what conceptual components students were emphasizing. Xue and Stains describe one analogy that has been used in instruction in which a dragon and a unicorn serve as resonance contributors with a rhinoceros as the resonance hybrid (hereafter referred to as the rhinoceros analogy).[4] The rhinoceros analogy and the picture analogy are similar in that they capture the idea of resonance hybrids incorporating the characteristics of resonance contributors. Both analogies also account for the fact that no singular resonance contributor fully represents the structure of a molecule. However, these two points are not as clearly illustrated by the teamwork analogy, which focuses on electron delocalization rather than the representational limitations of Lewis structures. A weakness of both the picture and teamwork analogies is also evident through comparison to the rhinoceros analogy. In the rhinoceros analogy, the two resonance contributors (the dragon and unicorn) do not exist in real life, accounting for the fact that the resonance contributors do not fully represent a molecule at any point in time. However, the picture and teamwork analogies do not account for this, as the resonance contributors in both analogies (each individual photo and each configuration of the team, respectively) represent the actual object in some way in which the object appears in life.

Through this comparison, the analogies students used suggest the conceptual components of resonance they may find more fundamental. Students using the picture analogy might be more focused on the idea that individual resonance contributors do not provide a complete structural representation, while students using the teamwork analogy might be more focused on the delocalization of electrons. However, it is pertinent to note that the WTL assignment indicated that "the primary goal of your writing is to teach the idea that no single contributing resonance structure represents the actual molecule," which is achieved by the picture analogy. Nevertheless, the commonly used student analogies do not reflect the idea that individual resonance contributors do not exist by themselves, which is communicated by the rhinoceros analogy. The fact that the second-semester students commonly used analogies that did not account for this part of resonance might suggest that they hold the incorrect understanding that resonance contributors exist as distinct structures in real time, as identified in previous studies.[2,3,37,38] However, with this finding, it should be noted that creating an analogy outside of naturalistic settings is a challenging task that may account for the weaknesses within students' commonly used analogies.[39] Nevertheless, this finding contributes to the existing literature by indicating that students are able to conceive of analogies that reflect aspects of their conceptualization of resonance and how resonance is represented. This finding points to the value in engaging students in creating analogies as a meaning-making task for supporting students' conceptual engagement within the framework of WTL.

Students primarily used carbonyl-containing compounds instead of benzene to explain resonance. Students were asked to use a specific example to explain how resonance influences reactivity, but there was no requirement to use an example molecule to explain resonance in general. However, 86.7% of students included an example molecule in their explanation. Hence, it is evident that many students found molecular examples to be useful for their discussions and explanations of resonance. Notably, 34.1% of students used a carboxylic acid or derivative that exhibited resonance. This was the most prevalent category of example molecules. One student included this example by writing,

> "Some atoms will be charged no matter the structures, but are able to use resonance to maintain their stability and low reactivity. An example of this is an acetate ion, there is a negative charge on one oxygen and a double bond on the other. Resonance allows the delocalized electrons on the negatively charge [*sic*] O to create a double bond and move the electrons from that double bond onto the O making it negative. This way neither O has to carry the burden of a full negative charge, but instead both possess a partial negative. This allows the acetate ion to remain more stable and less reactive."

This student selected a deprotonated carboxylic acid to explain how resonance can stabilize a charge by distributing it throughout a molecule. The presence of electronegative oxygen atoms serves to make carboxylic acids a good example for this application of resonance.

In contrast, 15.4% of students who included an example molecule used the Kekulé structure of benzene. This result is notable, as the literature describes how the nature of this structural representation can confuse novice students.[3,40] The fact that more students used a carboxylic acid derivative to show the delocalization of electrons between resonance contributors could indicate that students find this example clearer for communicating the concept. However, the assignment did give students the example of resonance contributors for the carbonate ion, and although only 8.8% of students who used an example molecule used the carbonate ion, this priming might have influenced students to include examples of similar carbonyl-containing compounds. Whether students find acyl compounds to better support their understanding of resonance compared to benzene would merit further research. Nevertheless, our finding suggests that students may value specific molecular examples over others for explaining resonance.

Students described resonance with an operational rather than conceptual definition. Our analysis also examined the definitions of resonance students shared with the general audience, which may indicate the features they found most helpful and important in their learning of the concept (Table 6.1). For this, 98.1% of students discussed that electrons are in different locations for different resonance contributors. One student described this by writing,

> "Basically, certain molecules have multiple diagrams because some electrons, known as delocalized electrons, are not committed to a single location meaning the diagram changes based on where you place electrons."

This student identified that, for molecules that exhibit resonance, the electrons are not fixed and can be drawn in different locations, resulting in different resonance structures. However, only 29.5% of students wrote that the atoms and connectivity of the molecule remain unchanged between resonance contributors. Furthermore, only 15.2% of students discussed when resonance can and cannot occur or what structural features allow for a given molecule to have resonance, a feature highlighted in the analytical framework's first LO.[6] This data suggests that many responses did not address some conceptual aspects of resonance that would be expected when explaining the concept and instead focused on the operational definitions for how to draw resonance structures. In other words, the second-semester students highlighted aspects of resonance they would be expected to demonstrate on typical assessments, such as being able to draw curved arrows to indicate the delocalization of electrons between resonance contributors, rather than the more conceptual aspects of resonance.

Table 6.1 Aspects of the resonance concept students included in their response

Component of resonance concept	Percentage	Exemplar
Electrons are in different locations in different resonance structures	98.1%	"Delocalizable electrons are electrons that can participate in resonance by moving from double bonds or lone electron pairs, to form new double bonds or lone pairs elsewhere on the molecule."
Atoms and connectivity are unchanged between resonance contributors	29.5%	"Resonance in chemistry terms is the idea that bonds in a molecule are able to move around and form multiple different structures while still maintaining the same connectivity (i.e. each atom is still connected to the same atoms, albeit with slightly different bonds)."
When resonance can/cannot occur	15.2%	"Now if you have enough molecules [*sic*] with p orbitals, the p-orbitals and all the electrons in it form a sort of continuous 'electron cloud' around the molecule. Now all the electrons in the p orbital are free to move around as they please and this causes the electrons to have more than one fixed arrangement. A molecule having different electron configurations is called resonance."

The choice of many students to not discuss the structural requirements for resonance is particularly important in light of Carle and Flynn's inclusion of being able to identify when resonance can occur as the first of their essential LOs.[6] The authors stress the importance of this LO, stating that identifying when resonance can occur "is a fundamental skill that is nested within many of the other learning outcomes" and that it "proves useful beyond organic chemistry, in disciplines such as biochemistry and physical chemistry."[6] However, the fact that only 15.2% of students included this LO in their explanation of resonance after engaging with the cognitive processes of writing suggests students might struggle with this conceptual aspect of resonance or did not find it an important aspect for their explanations. This finding contributes to Carle and Flynn's research by suggesting that, even in their second semester of organic chemistry, students are not necessarily considering the structural features allowing resonance to occur as a vital part of explaining resonance.

While students' responses largely did not consider when resonance can occur, their focus on the different configurations of electrons for each resonance structure indicates that students described resonance with a more operational definition reflective of how to draw resonance structures. Furthermore, students' use of language suggestive of electron movement aligns with the challenge in teaching resonance with the curved arrows that also depict electron movement within a reaction mechanism, as previously noted by Carle and Flynn.[6] Students' tendency to focus on this operational aspect of resonance is important to note, as prior studies have indicated that an operational understanding of resonance is not associated with students' conceptual understanding.[4,5] Furthermore, this finding speaks to the more general trend of students' rote understanding of concepts in organic chemistry focused on being able to answer exam questions instead of fully understanding concepts.[41–45]

6.7.2 How do second-semester organic chemistry students explain how resonance influences reactivity?

Students infrequently applied their understanding of resonance to a specific application when describing the influence of resonance on reactivity. While the prompt directed students to include a specific example of the influence of resonance on reactivity, 78.9% of students mentioned reactivity in a more general sense (as described in more detail below). Few students incorporated a specific application, and those who did described the application of resonance to acid–base chemistry, which was present in 12.2% of all responses. These students specifically discussed the role resonance plays when considering acid strength, explaining how resonance-stabilized conjugate bases correspond to more acidic acids than nonresonance-stabilized conjugate bases. One student wrote,

> "To tie this all together and mention one last important point about resonance and reactivity, lets [*sic*] consider two acidic molecules: one with resonance, and one without. Acids give up their hydrogens. When the H's leave, two electrons are left behind. However, the molecule without resonance cannot do anything to move these electrons around and the negative charge is 'stuck' in one place. In the resonating acid, the electrons can move around to different, more favorable positions. Thus, this negative charge is 'shared' throughout the molecule. Thus, this molecule is more stable than an acid where the negative charge is stuck on one, unfortunate atom."

This student highlights the effect that resonance stabilization has on relative acidity. By comparing two hypothetical acids, one with the potential to distribute the negative charge in the conjugate base, the student clearly articulates the role of resonance in acid–base chemistry. This result aligns with prior studies focused on students' understanding of acid–base chemistry by suggesting that some students can successfully explain the influence of resonance on reactivity using its application for determining relative acidity.[38,46] The finding contributes to these prior studies by indicating that the relationship between resonance and acidity is the most common specific application students used in their own explanations on the influence of resonance on reactivity.

This finding suggests that Carle and Flynn's eighth LO (using delocalization concepts to rank relative acidity/basicity) has been met by this subset of students at this point in their second semester introductory organic chemistry courses. However, Carle and Flynn's sixth and ninth LOs (determining whether a molecule is aromatic and determining and justifying the electrophilic and nucleophilic sites of a molecule using delocalization, respectively) were not addressed through specific examples in students' work. This is notable, as these LOs correspond to essential skills and knowledge related to resonance at the introductory organic chemistry level.[6] Hence, the fact that students did not include these aspects of resonance through a specific example suggests that they did not find them important applications of the concept for their audience during the cognitive processes of writing. However, it is possible that students did not choose these applications due to the nature of the assignment targeted toward a general audience. Nevertheless, this finding contributes to the literature on students' understanding of resonance by suggesting that the students who do consider a specific application for explaining the influence of resonance on reactivity tend to do so in the context of determining relative acidity.

Students equated resonance with stability but tended to have less nuanced explanations of how resonance influenced reactivity. A common trend throughout students' responses was to associate resonance with an increase in stability, with 81.0% of students stating that resonance would lead to a more stable molecule. Students provided various reasons to support this claim, with the common theme that the ability to distribute a negative charge throughout a molecule makes it more stable. For example, one student wrote,

> "When a molecule has resonance, the different atoms making up the molecules can share this charge, since their electrons can move throughout the molecule more easily. Sharing the charge makes molecules more stable."

This claim relates to Carle and Flynn's seventh LO (using delocalization concepts to determine and justify the relative stability, energy, and reactivity of ions) but only refers to resonance in general rather than using resonance to compare the relative stabilities of multiple ions.[6] This finding suggests that students recognize the role of resonance in distributing charges across a molecule to increase stability. However, students' association of the word "resonance" with "stability," without considering the context of comparing multiple structures, suggests that the second-semester students may hold a rule-based understanding of the relationship between resonance and stability. This finding contributes to the existing research suggesting students' rule-based understandings of other concepts within organic chemistry.[41−45]

Beyond associating resonance with stability, students also discussed the effect of stability on reactivity in a general sense, rather than discussing reactivity within the context of a specific application. Some students (27.1%) indicated that resonance stability would lead to decreased reactivity. By doing so, these students may have disregarded the fact that resonance stabilization does not inherently mean a molecule is unreactive (e.g., many reactions in introductory organic chemistry courses include starting materials that exhibit resonance, such as acyl transfer or electrophilic aromatic substitution reactions). Other students (36.5%) stated that resonance stabilization would lead to increased reactivity. The students who made this claim did so in two ways: both within a specific application or as a more general statement (see Table 6.2). When using a specific application, students tended to focus on using resonance to explain relative acidity, as described above. The remaining students (15.3%) took a more situationally dependent approach to characterize the effect of resonance stabilization on reactivity, writing that both an increase and decrease in reactivity could happen, depending on the situation (Table 6.2). However, many of these students appeared to treat resonance contributors as structures that exist as separate entities with different reactivities, as suggested by the exemplar response in Table 6.2. This finding provides further support to the previously published conceptual misunderstanding that resonance contributors are entirely different compounds that exist in equilibrium with one another.[1−3,37,38]

Table 6.2 Students' descriptions of the influence of resonance on reactivity.

Influence of resonance on reactivity	**Percentage[a]**	**Exemplar**
Lowers reactivity	27.1%	"Because resonance causes stability, we would say that resonance stabilized molecules are less reactive than those that aren't… Molecules that are stable are happy in their stability and thus, wouldn't want to change. So, if they don't want to change, they aren't going to want to react, and thus are less reactive molecules."
Increases reactivity (with a	36.5%	"An example of a resonance stabilized molecule would be the conjugate base of carboxylic acid. If you were to look at carboxylic acid... you would see the two possible resonance structures which are formed as the electrons within

specific application)		the double bond move around the compound. This ability of the electrons to move around cause [*sic*] this to be a highly reactive acid, meaning it is likely to lose its hydrogen atom."
Increases reactivity (in general)		"Structures which have resonance can be more reactive, because they have resonance, and are therefore more flexible with where their electrons go."
Both lowers and increases reactivity	15.3%	"For example, when electrons are shared more with an oxygen molecule, the oxygen becomes more negatively charged and, in turn, a much more reactive atom because it can use those extra electrons to make connections with other atoms. On the other hand, when the electrons are shared less with oxygen and more with another atom, the oxygen becomes more neutral and therefore a less reactive atom."
[a] The percentages are out of students who indicated that resonance increases stability, the remaining 21.1% of students who indicated that resonance increases stability did not say that this stability would increase or decrease reactivity.		

This set of findings highlights the trend for some students to overgeneralize the role of resonance on reactivity by making broad statements that resonance stability leads to lower and/or higher reactivity. Students' overgeneralization of how resonance stability influences reactivity, with their varied interpretations, provides further evidence for students' rule-based understanding of concepts in organic chemistry while highlighting the limitations of this narrow conceptualization.[41−45] Specifically, students' associations of stability with different interpretations of how stability influences reactivity indicates that students may not hold a nuanced understanding of how resonance influences reactivity. This finding contributes to the literature on students' understanding of resonance in general by identifying that second-semester students can hold diverging and overgeneralized viewpoints with regard to how resonance influences reactivity. Students' diverging viewpoints of how stability influences reactivity point to the need for further research and improved instruction on the topic to better support students' understanding of how resonance influences reactivity.

6.8 Limitations

There are key limitations associated with this study. Foremost is that this is a qualitative study conducted at a large, research-intensive institution in the United States. The findings may have limited transferability to other contexts, particularly those with differences in the introductory organic chemistry course sequence. Another limitation is that the writing assignment did not appear to elicit students' writing in alignment with four of the LOs in the analytical framework. While this may suggest a gap in students' understanding, this could also be due to a limitation of the assignment design (such as the intended audience for students' responses). Further research would be necessary to determine the influence of the prompt on students' responses and to elucidate students' understanding in alignment with the LOs not elicited by the prompt in this study. Furthermore, there are limitations associated with the WTL assignment design. Specifically, students that are multilingual writers or English language learners (ELL) may face additional challenges with describing resonance to a general audience in their responses. Future research is merited for how ELL students respond to similar WTL assignments in STEM; additionally, instructors should consider providing support to ELL students if choosing to implement similar WTL assignments in their courses (such as the peer review process and writing fellows implemented in this study). A final limitation regards the absence of interview or observational data collected for this study. As interviews with students were not conducted, there are limitations regarding the extent of the claims that can be made about students' choices in their writing. As instructional observations were not conducted, we are unable to make claims about where students may have drawn their explanations of resonance from (such as the analogies and examples students used).

6.9 Conclusions and implications

In this study, we used second-semester organic chemistry students' responses to a WTL assignment asking them to introduce the concept of resonance to the public to understand how students conceptualize resonance and its influence on reactivity. This study contributes to the literature by analyzing a large sample of student's written responses to an open-ended writing assignment as a means to gain insight into their understanding of the concept. Furthermore, by focusing on second-semester students, this work identifies the components of resonance that students found fundamental for their explanations after they had experience with the concept throughout their first semester of organic chemistry. By contributing to the literature through analysis of second-semester students' responses to an open-ended prompt about both resonance and how it influences reactivity, our findings can be used to influence instruction to support students' learning.

Our results suggest that students were able to use analogies to describe the concept of resonance and how resonance is represented. However, the analogies students used highlight some possible conceptual misunderstandings students may hold. Specifically, no student analogies accounted for the fact that resonance contributors do not exist in equilibrium with each other in real time. This is an important detail for instructors to convey when teaching resonance. Further research is warranted to better understand how students developed and selected the analogies they included in their writing and how their writing may have been shaped by the motivational, affective, and social components of WTL. Our findings also suggest that the Kekulé representation of benzene may not be as useful for students' understanding as carboxylic acid derivatives, which they used more commonly to explain resonance in their writing. Further research should seek to determine whether this finding reflects how students learn best, which could influence instruction in terms of what molecular examples to use when explaining resonance. Our findings regarding students' analogies and examples indicate how students themselves choose to explain the resonance concept; these findings suggest that instructors should carefully select a range of analogies and examples when introducing and teaching the concept. By using different analogies, for example, instructors could emphasize different aspects of resonance and how it is represented.

Students also commonly emphasized an operational rather than conceptual definition of resonance in their written explanations. Instead of discussing fundamental features of the concept, such as when resonance can occur and the fact that the atomic composition and connectivity remains unchanged between resonance contributors, students focused on how electrons are moved when drawing resonance structures. This finding suggests that students in their second semester of organic chemistry instruction may not give importance to certain fundamental features of resonance. To ensure that students have a true conceptual understanding of this topic instead of an operational understanding necessary to draw resonance structures, the methods of instruction and assessment for this concept may need to be revised to highlight the fundamental conceptual features.

When describing the influence of resonance on reactivity, students seldom provided specific applications or contexts for considering reactivity. Further research would be necessary to determine the roles of both the prompt and the curriculum that may have contributed to students' decisions not to discuss reactivity in more nuanced and complex ways. Instead of describing specific situations, it was common for students to equate resonance, in general, with molecular stability. However, students did not develop nuanced connections between how molecular stability from resonance would influence reactivity. Rather, students presented diverging and typically overgeneralized interpretations of how resonance stabilization can lead to either increased or decreased reactivity. This result contributes to the existing literature focused on how students understand resonance by demonstrating the varied and diverging ways that students interpret the influence of resonance on reactivity. These findings suggest that further research needs to be conducted to more deeply investigate how students conceptualize resonance and its influence on reactivity. Additionally, more instructional focus needs to be placed on discussing how resonance can influence the reactivity of molecules within introductory organic chemistry courses. Specifically, instructors should emphasize the nuanced and situationally dependent relationships between resonance, stability, and reactivity.

6.10 Appendices

6.10.1 Appendix 1. WTL assignment

Translating Resonance: An Explainer

Consider structures A, B, and C shown below that represent resonance contributors for the carbonate ion (Figure 6.2). The three structures are not detectable by any experimental method used to examine the structure of carbonate, which indicates that they do not sufficiently describe the carbonate ion.[a] This example demonstrates the limitation of using Lewis structures to depict resonance because none of the structures represent the actual molecule, which is a hybrid of A, B, and C.

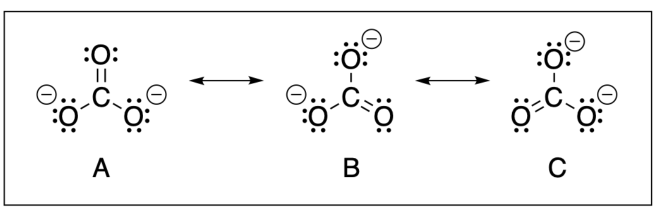


Figure 6.2 Structures that represent resonance contributors for the carbonate ion.

In his book, *Thing Explainer: Complicated Stuff in Simple Words*, Randall Munroe, who created the popular web comic series xkcd, challenges his readers to use the most common 1000 words in the English language to explain highly technical concepts. Neil Degrasse Tyson, Carl Sagan, and Bill Nye are well known for these types of translation skills – their ability to conduct rigorous scientific research and talk about why it matters helped make STEM accessible across generations of public audiences. Your objective for this assignment is to write an "explainer" piece that helps a general public understand *resonance*. An "explainer" is a brief radio spot or short journalistic piece, that aims to help a general public access, understand, and potentially apply a complicated or highly technical concept. The explainer often uses an entertaining, simple analogy to make an abstract, complicated idea simpler to understand and visualize. Explainers also give the reader a sense of why a concept is important for them to understand – what can or should we, the general public, do with the information that's been presented to us?

Here are a few sample explainers that cover concepts in other fields:

- "What the heck does ETF stand for?" from NPR's Marketplace:
  https://www.marketplace.org/2014/05/27/business/whiteboard%E2%84%A2-video/what-heck-does-etf-stand-explainer
- "Antibodies: Part I, CRISPR", a longer podcast from Radiolab explaining CRISPR, but you can find some great ideas in there, including the "repeat, blurb, repeat" spot on E-coli genetic code from 2:45-4:10:
  http://www.radiolab.org/story/antibodies-part-1-crispr/
- "Explaining supervolcanoes: big, hot, and dangerous", a monthly explainer from *The Guardian* in interview format:
  https://www.theguardian.com/world/2014/jan/09/supervolcanoes-yellowstone-bill mcguire

[a] http://www.chem.ucla.edu/~harding/tutorials/resonance/draw_res_str.html

Your explanation should help readers understand and interpret drawings of resonance contributors, their limitations, and how the contributors relate to the actual form of the molecule they represent.

Items to keep in mind:

- Your explainer should be between 350-500 words.
- Although the primary goal of your writing is to teach the idea that no single contributing resonance structure represents the actual molecule, within your explanation you should also make clear the implications of resonance for molecular structure of molecules.
- Provide and discuss a *specific* example in which resonance has implications on *chemical reactivity*. To identify possible examples for your explainer, you might want to consider what you know about resonance in class, the textbook.
- When you introduce technical terms, make sure you use simpler words to help define and translate them. Your audience is a general public listening to a science podcast or radio

program; they have not taken organic chemistry and are trying to understand resonance for the first time. Therefore, you should think carefully about the example(s) you present, whether the example(s) accurately portrays the resonance phenomenon, and whether it will help your readers to more easily understand it. Your readers look to you as an authority and you should carefully edit and proofread your essay so as to maintain credibility with them.

- Be creative – you could either write the explainer as if you are a scientist speaking directly to your audience or you can pretend that you're writing an interview between a journalist and a scientist; you can add short banter to the piece to get the audience interested; you can vary your punctuation to show emphasis or enthusiasm that you would want to come through if it were spoken aloud.
- External references are not required, but if they are used they should be cited using MLA format.

6.10.2 Appendix 2. WTL peer review criteria

- Would this explainer be understandable to someone from the general public who is unfamiliar with the concept of resonance? What is described well? What parts are difficult to understand?
- There should be an explanation on how resonance contributors relate to the actual form of the molecule. What is unclear? How can it be improved?
- The explanation should include the implications of resonance for molecular structure of molecules. What is missing or unclear? What is clear?
- There should be a discussion on the implications of resonance on chemical reactivity. Is a specific resonance example provided? Is the example accurate in portraying the resonance phenomenon? What is missing or unclear? What is clear? What can be covered in more detail?

6.10.3 Appendix 3. Coding scheme

Table 6.3 Coding scheme

Category	**Code (Learning Objective**	**Description**	**Exemplar**	**Percentage of responses**	**Krippendorff's alpha**

	Alignment)[a]				
Examples for explaining resonance	Analogy (N/A)	Student compares resonance to a common object, theme, etc. (When coding, include a brief description of the analogy.)	"We can have a hundred pictures of ourselves, each photograph revealing a different angle." "However, no single photograph carries all the information about us; we are essentially a hybrid of all the photographs."	74.3%	0.74[b]
	Example (N/A)	Student highlights a specific molecule to explain their description of resonance.	"We see this in molecules like ozone, where three oxygen molecules share electrons, and there are different ways the charge can be distributed."	86.7%	1.0[c]
Resonance definitions	Limitations of a single Lewis structure (N/A)	Student states that a single Lewis structure does not account for resonance (or dynamic nature/movement of electrons). Can apply even if this is a part of the hybrid structure definition. A single contributor or structure does not represent the true molecule.	"Each Lewis structure, which is a drawing of an arrangement of a molecule, carries some information about the molecule, but no single Lewis structure has the full story"	84.8%	1.0[c]
	Resonance hybrid (LO4)	Student discusses that a hybrid exists which is a combination of the resonance contributors. This is the most "accurate" representation of the structure.	"The molecule exists as a hybrid of each Lewis structure"	73.3%	0.70[b]
	Major and minor resonance contributors (LO3)	The student identifies that certain resonance contributors can be major and minor contributors. The student may identify what makes a contributor	"Major contributors are major because those structures are the most stable. Minor contributors	21.0%	0.77[b]

		major vs minor (closed shell atoms, uncharged, etc.). Student may describe the major contributor as the "most stable" resonance contributor.	are less stable, but could be very important for a molecule's reactivity."		
	Atoms in the same location / same connectivity (N/A)	Student identifies that between resonance contributors, the atoms and connectivity will remain the same. Student needs to be saying that between resonance structures, the connectivity stays the same (more than just defining what a Lewis structure is). The student can also say that the chemical formula stays the same.	"Each atom, and all the single bonds that connect the atoms between each other, are the same in each Lewis structure"	29.5%	1.0[c]
	Electrons in different locations (N/A)	Student identifies that between resonance contributors, the location of delocalized electrons will differ. Student may provide a definition here of what delocalization is and why some electrons are localized and some are delocalized.	"Delocalized electrons can exist as a second or third bond between two atoms. They can also exist as free floating electrons that simply float on an atom. The position of delocalized electrons is what changes between each Lewis structure."	98.1%	1.0[c]
	When resonance can occur (LO1)	Student explains what structural elements allow for resonance to occur or provide an example of when resonance can/cannot occur. Only applies when a student indicates that resonance only applies to certain molecules, such as by providing an example where resonance does not apply. Can just include specific details such as when resonance can or	"If we think about water, or H20 [*sic*], it's extremely stable. It does not like to react with other molecules. H20 [*sic*] has no resonance; it's electrons can't be rearranged."	15.2%	0.62[d]

		can not occur. Different than just highlighting the presence of delocalized electrons.			
Influence on reactivity	Resonance leads to more stability (LO7)	Student states that resonance allows for more stability (can give multiple reasons such as allowing electron density to spread over larger space, lowering energy). Should be saying that because of resonance the entire molecule is stable (e.g., not that charges are stabilized).	"Well, when a structure is stable, it's happy as it is. It does not want to react with other molecules because reacting with another molecule could make it unstable."	81.0%	1.0[c]
	Resonance indicates distribution of partial charges (LO7)	Student identifies that minor resonance contributors with charges can help to identify partial charges. This also has an impact on reactivity. Code applies to any reference the student makes about the creation or position of charges in resonance contributors.	"The second structure displays the positive 'character' of the carbon atom, demonstrating why it was attacked."	73.3%	1.0[c]
	Resonance stabilized conjugate bases are associated with stronger acids (LO8)	Student identifies that acids with resonance stabilized conjugate bases are more acidic, as the more stable base makes the loss of an acidic proton more favorable.	"The acid is a stronger acid because it is stabilized due to resonance in its conjugate base form."	12.4%	1.0[c]
	Resonance leads to increased reactivity (LO9)	Student argues that the presence of resonance will increase the reactivity of the molecule. Can also be coded for if the student argues for lower reactivity elsewhere in the paper. Include situations where they say because of resonance a reaction or some type of reactivity is possible. Examples of how charges interact don't necessarily count for this code.	"Structures which have resonance can be more reactive, because they have resonance, and are therefore more flexible with where their electrons go."	54.3%	0.90[c]

	Resonance leads to lower reactivity (N/A)	Student argues that the molecules stabilized by resonance are less reactive overall. Can also be coded for if the student argues for increased reactivity elsewhere in the paper.	"The ability of electrons to "delocalize" over 2 or more atoms stabilizes the molecule causing the product to be favored more often and makes the molecule less reactive."	35.2%	1.0[c]

[a] Alignment of individual codes with the learning objectives from Carle and Flynn[6] are indicated in parenthesis.
[b] Krippendorff's alpha value that is acceptable for tentative conclusions.
[c] Krippendorff's alpha value that is reliable.
[d] Krippendorff's alpha value that is below the cutoff for values that are acceptable for tentative conclusions.

6.10.4 Appendix 4. Coding scheme alignment with the analytical framework

The coding scheme was developed in alignment with the Learning Objectives (LOs) described in Carle and Flynn's "Essential learning outcomes for delocalization (resonance) concepts: How are they taught, practiced, and assessed in organic chemistry?"[6] LOs one, three, and four aligned with our first research question by corresponding to how students described the concept of resonance. The first LO focused on the ability of students to identify structures in which delocalization can occur.[6] This LO was reflected in the "when resonance can occur" code, highlighting specific instances where students described the structural elements that allowed resonance to be present. The third LO assessed the ability of students to identify which resonance structures would contribute more to the overall resonance hybrid. This LO is reflected in our "major and minor resonance contributors" code which highlights the ability of students to identify that certain resonance structures can be described as "major" and "minor" due to their contribution to the overall resonance hybrid. The final LO corresponding to students' descriptions of the resonance concept, the fourth LO in Carle and Flynn's framework, involved assessing whether students could draw the resonance hybrid and explain what it represented. Our "resonance hybrid" code aligns with this LO as it was applied to instances where students describe what the resonance hybrid represents.

LOs 7, 8, and 9 aligned with our second research question, examining how students explained the influence of resonance on reactivity. Carle and Flynn's seventh LO focuses on determining the stability, energy, and reactivity of ions. Our "resonance leads to more stability" and "resonance indicates distribution of partial charges" codes incorporate ideas from this LO in slightly different ways. These codes were applied to instances of students describing how resonance can stabilize structures in general and how charges can be formed or stabilized in resonance contributors, respectively. The "resonance stabilized conjugate bases are associated with stronger acids" code corresponded to the eighth LO, which dealt with students using delocalization concepts to explain relative acidity and basicity. This code was applied to examples where students described the effect of resonance or delocalization on relative acidities of molecules. The ninth LO involved students determining electrophilic and nucleophilic sites on molecules to predict reactivity. This aligned with our "resonance leads to increased reactivity" code, which included all examples of students describing how aspects of resonance can indicate how a molecule will react or increase overall reactivity in general.

The remaining LOs in Carle and Flynn's framework (LOs two, five, six, and ten) were not applicable to our coding scheme. The second LO involved students being able to draw resonance structures using curved arrows to show the delocalization of electrons. We did not code for this LO as the assignment did not involve students drawing resonance structures but rather describing them with words. The fifth and sixth LOs, which involved hybridization and aromaticity, were aspects of the resonance concept beyond the scope of an introductory explanation of resonance to the public, so no codes correspond to these LOs. Finally, we did not create a code for the tenth LO, which corresponds to students using delocalization concepts to predict and justify the outcome of a reaction, as the assignment did not involve students considering reaction schemes. Furthermore, there were no student responses aligned with LOs five, six, or ten within the coded data.

Our coding scheme also included codes that were not drawn from Carle and Flynn's LOs but were incorporated through open coding during the initial analysis stages. Two of these codes related to recording the analogies and examples that students used to explain resonance. The use of analogies was suggested and highlighted in the prompt of the assignment. As such, an "analogy" code was applied to instances where students included non-scientific examples to better convey the concept of resonance to their audience. The "example" code was also in this general coding category. This code was applied to situations where students included an example molecule within their response which they used to demonstrate or further explain resonance, beyond the examples students used to describe the influence of resonance on reactivity. We additionally noted the prevalence of different molecular examples students used. Additional codes that did not directly relate to Carle and Flynn's LOs included those that fell into the coding categories relating to resonance definitions and the influence of resonance on reactivity. These codes captured instances where students described the limitations of a single Lewis structure, stated that the atoms of a molecule remain in the same place between resonance contributors, described the delocalization of electrons between resonance contributors, and indicated that resonance lowers the reactivity of a molecule.

6.11 Acknowledgements

The authors would like to thank the Keck Foundation and the University of Michigan Third Century Initiative for funding. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260. The authors would additionally like to thank the students who participated in this study; Carol Ann Castaneda, the course instructor; Anne Ruggles Gere, Barry C. Thompson, Raymond Pugh, Ashley Karlin, and Atia Sattar who contributed to the WTL assignment development; and Rebecca Fantone, Solaire Finkenstaedt-Quinn, Michael Petterson, and other members of the Shultz group for their discussions during the preparation of this manuscript.

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

(3) Kim, T.; 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.

(4) 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.

(5) 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.

(6) 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.

(7) Rivard, L. O. P. A Review of Writing to Learn in Science: Implications for Practice and Research. *J. Res. Sci. Teach.* **1994**, *31* (9), 969–983. https://doi.org/10.1002/tea.3660310910.

(8) Reynolds, J. A.; Thaiss, C.; Katkin, W.; Thompson, R. J. Writing-to-Learn in Undergraduate Science Education: A Community-Based, Conceptually Driven Approach. *CBE Life Sci. Educ.* **2012**, *11* (1), 17–25. https://doi.org/10.1187/cbe.11-08-0064.

(9) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Res. Teach. English* **2015**, *50* (2), 199–235.

(10) Klein, P. D.; Boscolo, P. Trends in Research on Writing as a Learning Activity. *J. Writ. Res.* **2016**, *7* (3), 311–351. https://doi.org/10.17239/jowr-2016.07.03.01.

(11) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and Conceptual Learning in Science: An Analysis of Assignments. *Writ. Commun.* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(12) Finkenstaedt-Quinn, S. A.; Petterson, M.; Gere, A.; Shultz, G. Praxis of Writing-to-Learn: A Model for the Design and Propagation of Writing-to-Learn in STEM. *J. Chem. Educ.* **2021**, *98* (5), 1548–1555. https://doi.org/10.1021/acs.jchemed.0c01482.

(13) Shultz, G. V.; Gere, A. R. Writing-to-Learn the Nature of Science in the Context of the Lewis Dot Structure Model. *J. Chem. Educ.* **2015**, *92* (8), 1325–1329. https://doi.org/10.1021/acs.jchemed.5b00064.

(14) Moon, A.; Zotos, E.; Finkenstaedt-Quinn, S.; Gere, A. R.; Shultz, G. Investigation of the Role of Writing-to-Learn in Promoting Student Understanding of Light-Matter Interactions. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 807–818. https://doi.org/10.1039/c8rp00090e.

(15) Moon, A.; Moeller, R.; Gere, A. R.; Shultz, G. V. Application and Testing of a Framework for Characterizing the Quality of Scientific Reasoning in Chemistry Students' Writing on Ocean Acidification. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 484–494. https://doi.org/10.1039/c9rp00005d.

(16) Finkenstaedt-Quinn, S. A.; Snyder-White, E. P.; Connor, M. C.; Gere, A. R.; Shultz, G. V. Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *J. Chem. Educ.* **2019**, *96* (2), 227–237. https://doi.org/10.1021/acs.jchemed.8b00711.

(17) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Kasner, G.; Wilhelm, C. A.; Moon, A.; Gere, A. R.; Shultz, G. V. Capturing Student Conceptions of Thermodynamics and Kinetics Using Writing. *Chem. Educ. Res. Pract.* **2020**, *21*, 922–939.

https://doi.org/10.1039/C9RP00292H.

(18) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(19) Watts, F. M.; Schmidt-McCormack, J.; Wilhelm, C.; Karlin, A.; Sattar, A.; Thompson, B.; Gere, A. R.; Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21*, 1148–1172. https://doi.org/10.1039/C9RP00185A.

(20) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Chambers, T. G.; Moon, A.; Goldman, R. S.; Gere, A. R.; Shultz, G. V. Investigation of the Influence of a Writing-To-Learn Assignment on Student Understanding of Polymer Properties. *J. Chem. Educ.* **2017**, *94* (11), 1610–1617. https://doi.org/10.1021/acs.jchemed.7b00363.

(21) Halim, A. S.; Finkenstaedt-Quinn, S. A.; Olsen, L. J.; Gere, A. R.; Shultz, G. V. Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review. *CBE Life Sci. Educ.* **2018**, *17* (2), 1–12. https://doi.org/10.1187/cbe.17-10-0212.

(22) Gupte, T.; Watts, F. M.; Schmidt-McCormack, J. A.; Zaimi, I.; Gere, A. R.; Shultz, G. V. Students' Meaningful Learning Experiences from Participating in Organic Chemistry Writing-to-Learn Activities. *Chem. Educ. Res. Pract.* **2021**, *22*, 396–414. https://doi.org/10.1039/d0rp00266f.

(23) Grimberg, B. I.; Hand, B. Cognitive Pathways: Analysis of Students' Written Texts for Science Understanding. *Int. J. Sci. Educ.* **2009**, *31* (4), 503–521. https://doi.org/10.1080/09500690701704805.

(24) Arnold, K. M.; Umanath, S.; Thio, K.; Reilly, W. B.; McDaniel, M. A.; Marsh, E. J. Understanding the Cognitive Processes Involved in Writing to Learn. *J. Exp. Psychol. Appl.* **2017**, *23* (2), 115–127. https://doi.org/10.1037/xap0000119.

(25) Emig, J. Writing as a Mode of Learning. *Coll. Compos. Commun.* **1977**, *28* (2), 122–128.

(26) Klein, P. D. Reopening Inquiry into Cognitive Processes in Writing-To-Learn. *Educ. Psychol. Rev.* **1999**, *11* (3), 203–270. https://doi.org/10.1023/A:1021913217147.

(27) Prior, P. A Sociocultural Theory of Writing. In *Handbook of Writing Research*; MacArthur, C., Graham, S., Fitzgerald, J., Eds.; Guilford, 2006; pp 54–66.

(28) Flower, L.; Hayes, J. R. Images, Plans, and Prose: The Representation of Meaning in Writing. *Writ. Commun.* **1984**, *1* (1), 120–160.

(29) Flower, L.; Hayes, J. R. A Cognitive Process Theory of Writing. *Coll. Compos. Commun.* **1981**, *32* (4), 365–387.

(30) Hayes, J. R. A New Framework for Understanding Cognition and Affect in Writing. In *The Science of Writing: Theories, Methods, Individual Differences, and Applications*; Levy, C. M., Ransdell, S., Eds.; Lawrence Erbaum Associates: Mahwah, New Jersey, 1996; pp 1–27.

(31) Finkenstaedt-Quinn, S. A.; Polakowski, N.; Gunderson, B.; Shultz, G. V.; Ruggles Gere, A. Utilizing Peer Review and Revision to Support the Development of Conceptual Knowledge Through Writing. *Writ. Commun.*

(32) Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101.

(33) Miles, M. B.; Huberman, A. M.; Saldana, J. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed.; Sage: Los Angeles, CA, 2014.

(34) Watts, F. M.; Finkenstaedt-Quinn, S. A. The Current State of Methods for Establishing Reliability in Qualitative Chemistry Education Research Articles. *Chem. Educ. Res. Pract.* **2021**, No. 22, 565–578. https://doi.org/10.1039/d1rp00007a.

(35) Kirilenko, A. P.; Stepchenkova, S. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLoS One* **2016**, *11* (3), 1–14. https://doi.org/10.1371/journal.pone.0149787.

(36) Krippendorff, K. Bivariate Agreement Coefficients for Reliability of Data. *Sociol. Methodol.* **1970**, *2*, 139–150.

(37) Finkenstaedt-Quinn, S. A.; Watts, F. M.; Petterson, M. N.; Archer, S. R.; Snyder-White, E. P.; Shultz, G. V. Exploring Student Thinking about Addition Reactions. *J. Chem. Educ.* **2020**, *97* (7), 1852–1862. https://doi.org/10.1021/acs.jchemed.0c00141.

(38) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.; Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions via App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 878–892. https://doi.org/10.1039/c9rp00260j.

(39) Dunbar, K. The Analogical Paradox: Why Analogy Is so Easy in Naturalistic Settings, Yet so Difficult in the Psychological Laboratory. In *The Analogical Mind: Perspectives from Cognitive Science*; Gentner, D., Holyoak, K. J., Kokinov, B. N., Eds.; 2001; pp 313–334. https://doi.org/10.1353/lan.2006.0085.

(40) Kozma, R.; Chin, E.; Russell, J.; Marx, N. The Roles of Representations and Tools in the Chemistry Laboratory and Their Implications for Chemistry Learning. *J. Learn. Sci.* **2000**, *9* (2), 105–143. https://doi.org/10.1207/s15327809jls0902_1.

(41) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable Reasoning: Multi-Variate Problem-Solving in Organic Chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281–292. https://doi.org/10.1039/c0rp90003f.

(42) Christian, K.; Talanquer, V. Modes of Reasoning in Self-Initiated Study Groups in Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 286–295. https://doi.org/10.1039/c2rp20010d.

(43) Grove, N. P.; Bretz, S. L. A Continuum of Learning: From Rote Memorization to Meaningful Learning in Organic Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 201–208. https://doi.org/10.1039/c1rp90069b.

(44) Cruz-Ramírez De Arellano, D.; Towns, M. H. Students' Understanding of Alkyl Halide Reactions in Undergraduate Organic Chemistry. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 501–515. https://doi.org/10.1039/c3rp00089c.

(45) Weinrich, M. L.; Talanquer, V. Mapping Students' Modes of Reasoning When Thinking about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 394–406. https://doi.org/10.1039/c5rp00208g.

(46) Cartrette, D. P.; Mayo, P. M. Students' Understanding of Acids/Bases in Organic Chemistry Contexts. *Chem. Educ. Res. Pract.* **2011**, *12* (1), 29–39. https://doi.org/10.1039/c1rp90005f.

Chapter 7

Investigating Writing-To-Learn To Elicit Organic Chemistry Students' Reasoning About Alternative Reaction Mechanisms

7.1 Initial remarks

This chapter is the second of three chapters which explore the analysis of student responses to different writing-to-learn (WTL) assignments to examine how WTL can elicit students' reasoning in organic chemistry. This chapter and Chapter 8 both explore how students reason with reaction mechanisms, which are a central feature of organic chemistry courses. These two chapters contribute to the research literature on reaction mechanisms in organic chemistry education by demonstrating how WTL assignments can both support and elicit students' reasoning with reaction mechanisms. This chapter specifically focuses on a WTL assignment that elicits students' considerations of alternative reaction pathways for a single reaction, focusing on students' reasoning for the relative likelihood of the two pathways when considering the two most prevalent representations for reaction mechanisms: the electron-pushing formalism and reaction coordinate diagrams.

The study in this chapter is guided by a framework for representational competence, which is the ability to use representations to describe and explain phenomena. One of the primary representational competence skills is the ability to recognize the appropriate representations for making and justifying specific claims. For the WTL assignment, the reaction coordinate diagram is the more appropriate representation for making a claim about relative likelihood of the two reaction pathways. The analysis of students' responses involved both qualitative and quantitative approaches to identify students' reasoning and how their reasoning changed after peer review and revision. The results revealed various features of the two representations that students used in their writing, both in their initial and revised drafts. Specifically, all students incorporated features from both representations in their responses; however, students making an incorrect claim about the relative likelihood of the two reaction pathways reasoned by appealing more often to the electron-pushing formalism compared to students who correctly selected the more likely pathway. The analysis of the peer review process involved identifying points of disagreement students encountered in the peer review comments they received and in the drafts they read. Using logistic regression, the analysis revealed that both components of the peer review process influenced students' revisions, though students' revisions to change their selection of the more likely mechanism were more influenced by reading drafts with alternative viewpoints (compared to receiving comments with alternative viewpoints). The findings from this study provide implications for teaching students to reason about reaction mechanisms using multiple representations while highlighting the utility of the peer review process for engaging students in revisiting their reasoning. Furthermore, the findings contribute to the literature by highlighting how students' mechanistic reasoning relates to representational competence.

This chapter was originally published as a research article in *Chemistry Education Research and Practice*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author, I contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing). G.Y. Park and M.N. Petterson, undergraduate mentees, contributed to data analysis. G.V. Shultz contributed to funding acquisition, project supervision, conceptualization, data collection, and writing (review and editing).

Original publication and copyright information:

Reproduced from F.M. Watts, G.Y. Park, M.N. Petterson, and G.V. Shultz, *Chem. Educ. Res. Pract.* 2022, **23**, 486–507 with permission from the Royal Society of Chemistry.

7.2 Abstract

Organic reaction mechanisms are often represented by the electron-pushing formalism and reaction coordinate diagrams. These representations pose a challenge to students because valuable information is encoded within each representation, and students must know how to reason about mechanisms using both. Hence, it is important to understand whether and how students consider these two representations when reasoning about reaction mechanisms. We have collected responses to a writing-to-learn assignment administered in a second-semester organic chemistry laboratory course to investigate students' reasoning. The assignment was designed to elicit

students' reasoning about the most likely of two mechanisms for a catalyzed intramolecular aldol reaction when given the electron-pushing scheme and reaction coordinate diagram for both mechanisms. As part of the assignment, students submitted initial drafts, participated in content-focused peer review, and submitted revised drafts. We analyzed each component using a mixed methods approach to identify students' reasoning about the most likely reaction pathway and how their reasoning changed after peer review and revision. In this article, we present a quantitative overview of changes students made about their decisions for the most likely reaction pathway and how these changes are related to providing and receiving feedback. Additionally, we present our analysis of the features of representations students used to reason about the likelihood of alternative reaction mechanisms. This study demonstrates how existing research about students' reasoning with representations was operationalized for classroom practice using writing-to-learn. Furthermore, the analysis illustrates how writing-to-learn to can be used to develop students' reasoning and offers implications for teaching students to reason about reaction mechanisms using multiple representations.

7.3 Introduction

Students typically encounter two representations of organic reaction mechanisms in introductory organic chemistry courses: the electron-pushing formalism (EPF) and reaction coordinate diagrams (RCDs). A growing body of research examines how students reason about organic reaction mechanisms with the EPF.[1] More recently, attention has focused on how students reason with RCDs,[2–4] with one study examining how students match reactions to RCDs.[5] However, few existing studies explore how introductory organic chemistry students reason with both of these representations together. As such, the goal of this study is to explore students' reasoning in writing when considering both representations of reaction mechanisms. We achieve this through a writing-to-learn (WTL) assignment that asked students to reason about the most likely of two mechanisms for a single transformation, given the EPF schemes and RCDs for both mechanisms. The WTL assignment was implemented with peer review and revision, which allowed us to further investigate (1) how students' reasoning with these representations changed during the weeks spanning the assignment and (2) how changes in students' reasoning from their initial to final drafts might be influenced by the peer review process. In this work, we present our analysis of students' reasoning as presented across their initial and final responses to the WTL assignment.

We additionally present our analysis of the influence of the peer review process. This research demonstrates the operationalization of chemistry education research findings to inform the design of WTL and to further investigate students' reasoning with mechanistic representations when they write about organic reactions.

7.3.1 Reasoning with mechanistic representations in organic chemistry

The Next Generation Science Standards identifies multiple scientific practices that science educators should focus on for improving STEM education. One of these scientific practices is "developing and using models."[6] In this context, a "model" is not only the representation that depicts a phenomenon but also the cognitive processes involved in developing and using said representation. In other words, models include the graphs, figures, and/or structures that scientists use to depict a phenomenon along with the epistemic practices for developing models and using them to explain or predict phenomena.[7,8] As such, researchers emphasize that instruction in alignment with this scientific practice should not only encompass what a model is *of*, but also what it is *for*.[9]

Models and representations are central to reasoning in organic chemistry, as evident by the nature of chemical knowledge spanning the submicroscopic, macroscopic, and symbolic domains described by Johnstone's triangle.[10,11] Representations lie within the symbolic domain and are used to represent and bridge between submicroscopic and macroscopic concepts.[10] There are multiple representations within the symbolic domain in the context of organic reaction mechanisms, including energy graphs and molecular structures.[11] The two symbolic representations of organic reaction mechanisms often taught in introductory organic chemistry courses are the EPF and RCDs. Both representations provide different information about reactions that reflect organic chemists' conceptions of chemical transformations.[12] As such, the learning goals for teaching these representations are for students to understand (1) how the representations align with chemical ideas or concepts and (2) how the representations can be used to construct claims, predictions, or explanations.[13–15]

Students' reasoning with the EPF has received significant attention in the literature. Many studies examine how students make connections between the EPF representation and the underlying chemical properties that guide the proper use of the EPF. These studies indicate that students often focus on surface features, such as charges, when reasoning through reaction mechanisms.[16–19] Students also have challenges making connections between structure and function when reasoning about reaction mechanisms using key concepts including resonance,[20–22] nucleophilicity,[16,19,23–26] and sterics.[27] Other studies similarly suggest that students focus more on the surface features and structures of the representation rather than the implicit chemical properties and functions.[23,26,28] This understanding of how students conceptualize the EPF provides a valuable basis for exploring how students use the EPF when reasoning, making claims, or constructing explanations.

Studies in the literature also examine students' understandings of RCDs with investigations into how students interpret the meaning of RCD surface features.[2–5,29–31] The studies by Popova and Bretz specifically report several findings about how organic chemistry students understand RCDs.[3–5] For example, Popova and Bretz identified that students in their study often viewed RCDs as encoding information that reflects only the major reacting species rather than all components of a reaction, often not considering the submicroscopic level.[3] In another article, Popova and Bretz identified that students demonstrated challenges when translating between mechanisms and RCDs for substitution and elimination reactions due to incomplete understandings of the information communicated by RCD surface features.[5] They also found students have difficulty with this task because their reasoning with the mechanisms was often product-oriented and focused on the surface features of reactants.[5] The Popova and Bretz articles were all in the context of a course that primarily taught RCDs alongside reactions in first-semester organic chemistry (e.g., substitution, elimination, and addition reactions) but not in the second semester of instruction.[3–5] Hence, their findings suggest the need to provide students with further opportunities to develop their reasoning with RCDs. Furthermore, across the articles describing students' interpretations of RCD surface features, researchers identified that students often conflate transition states with intermediates, do not note energy changes encoded on the *y*-axis, and view the *x*-axis as corresponding to time.[2,4,31,32] The findings regarding students' understandings of RCD surface features and how they connect to mechanisms suggest a need to further support students' use and understanding of RCDs within the organic chemistry curriculum. Furthermore, the findings provide a baseline for understanding how students might use these representations in their reasoning.

As the existing literature indicates, researchers are focused on how students connect the surface features of both representations to chemical ideas, properties, and concepts. However, further research is necessary to understand how students use these representations in their reasoning. Some existing studies provide insight into how students reason in chemistry.[33–37] Specifically, studies demonstrate that few students reason based on mental models that relate structure to reactivity, while many students rely on memorized rules or cases.[33,34] Similarly, students often face challenges with integrating multiple variables into their reasoning or using reasoning to connect evidence to claims.[35–37] The challenges students have with reasoning in organic chemistry specifically may be related to a tendency for rote memorization rather than meaningful learning.[38] More recent studies examined students' engagement with contrasting cases and how these types of problems can encourage students to consider multiple conceptual factors when producing an explanation.[18,27,39,40] Nevertheless, students in these studies still often exhibited limited complexity in their reasoning or explanations. Altogether, these studies provide evidence of students' abilities for both reasoning and providing explanations in organic chemistry across different problem types. These studies point to the need for further research into understanding how students reason and develop explanations when using two common mechanistic representations in organic chemistry. The goal of this study is to address this need through implementing a WTL assignment in the classroom that targets this aspect of students' reasoning.

7.3.2 Using writing-to-learn, peer review, and revision to access students' reasoning

Prior research demonstrates the analysis of students' writing to access their reasoning about STEM content in organic chemistry[19,41] and other content areas.[42–44] Some of these studies specifically elicited students' reasoning through WTL,[19,41,43] an instructional practice that emphasizes the role of writing assignments in supporting students' conceptual understanding.[45,46] Studies demonstrate that WTL is effective for supporting understanding in a variety of STEM courses, including chemistry, biology, materials science, and statistics.[47–53] In addition to the writing assignments supporting students' conceptual understanding, WTL also incorporates peer review and revision that provide further learning opportunities. This aspect of WTL pedagogy can also explain how students' reasoning might change for specific content due to these structures.[54] As these prior studies suggest, students' responses to the WTL process are a valuable source of data for accessing students' reasoning.

Existing studies of WTL in STEM courses examine the role of peer review in supporting students' conceptual understanding and the revisions students make.[48,51,53,55] Specifically, studies demonstrate that students can use the peer review process to provide content-focused, constructive feedback.[51,55] However, students do not always indicate incorrect content when commenting on other students' drafts.[51] In addition, while students tend to make revisions in general, their revisions do not always necessarily align with the peer review comments they received.[50,55] Furthermore, peer review and revision can serve to both remediate some misunderstandings while also surfacing additional misunderstandings that were not present in students' initial responses.[48] Altogether, these studies suggest that receiving peer review comments is valuable for encouraging revision but not necessarily for remediating students' incorrect understanding. However, studies do suggest that reading other students' work and providing feedback may have more influence on students' revisions compared to receiving feedback, in STEM courses,[53] writing courses,[56–58] and during hypothetical peer review.[59] In addition to supporting students' conceptual learning, there is also evidence that peer review can support students' positive affective experiences with WTL assignments in a way that supports meaningful learning.[60,61] With the existing evidence for the role of WTL with peer review and revision supporting students' learning, it is necessary to further explore WTL assignments and the peer review process in the context of students' reasoning with representations in organic chemistry, which is the goal of this study. Furthermore, using WTL assignments in this way demonstrates how the existing research findings pertaining to students' reasoning with representations and case comparisons can be operationalized and implemented within the classroom to further investigate students' reasoning.

7.4 Theoretical perspectives

This study is informed by the representational competence framework and the cognitive process theory of writing. Representational competence provides explanatory power for investigating students' reasoning with multiple representations, which we captured in their writing. The cognitive process theory of writing is a complementary perspective that we used to uncover students writing processes as they engaged with this WTL assignment.

7.4.1 Representational competence

As described by Kozma and Russell,[14] representational competence is the ability to use representations to describe and explain chemical phenomena; this aligns the idea that the scientific practice of "developing and using models" includes both the representation of a phenomenon and how the representation is used in practice.[7–9] Kozma and Russel's framework recognizes that much of chemistry instruction surrounds the various ways that chemists represent sub-microscopic phenomena. All representations in the context of chemistry require an understanding of the related chemical concepts, including the two representations central to this study, the EPF and RCDs. As described in the literature review, students' use of these representations in organic chemistry tends to emphasize surface features rather than underlying chemical concepts, which is a reflection of novice representational use.[13,14,62] However, another feature of representations that merits further study is that they are inherently required for communication, supporting claims, or making predictions.[13,14] For instance, the EPF is useful for explaining or predicting the chemical structure of reaction products; similarly, RCDs are useful for explaining the thermodynamic and kinetic parameters that control the products of a reaction.[63] Representational competence is also necessary for communicating concepts, ideas, or claims surrounding chemical phenomena.[13,14] Beyond being able to use representations to support claims or make predictions, representational competence also encompasses selecting the appropriate representation for making particular claims or predictions.[14] This latter ability reflects the idea that, in many cases, more than one representation can be used to describe the same chemical phenomena—and that two representations of the same phenomenon can provide both similar and unique information.[14] Because of this, it is important to investigate how students use multiple representations when supporting claims or making predictions in organic chemistry. Furthermore, it is necessary to extend the existing research literature investigating students' representational competence so instructors can better inform their practice of teaching, eliciting, and assessing representational competence in the classroom.[15]

7.4.2 Cognitive process theory of writing

The other theoretical framework guiding this study supports the utilization of WTL both within the classroom and for investigating students' reasoning. As described in previous studies using data from WTL assignments,[19,41] the cognitive process theory of writing underpins the analysis of written responses for accessing students' understanding.[64–66] The cognitive process theory suggests that students' written responses reflect the concepts and knowledge they used throughout the process of writing and revising in response to an assignment. Furthermore, the theory emphasizes writing as a recursive process in which planning, drafting, and revising occur throughout all stages of the writing process. As cognitive writing processes require producing internal representations of knowledge that engage both long- and short-term memory, the texts that students produce make visible the concepts used to respond to a writing task. The cognitive process theory also suggests the value of implementing writing assignments with peer review and revision to further support students' learning, as these structures provide further opportunity to engage with the cognitive writing processes. With the focus on recursive processes and revision, the cognitive process theory aligns with the model of cognition which suggests that the concepts and ideas used to respond to a task are activated within a specific context and can change across time.[67,68] Hence, cognitive process theory provides a lens through which to understand how writing about representations engages students in the aforementioned aspects of representational competence across the initial draft and revision components of WTL assignments.

7.5 Research questions

This study examines introductory organic chemistry students' reasoning when considering multiple representations of organic reaction mechanisms. Through our analysis of students' reasoning as presented in their responses to a WTL assignment, we seek to address the following research questions:

1. What features of multiple mechanistic representations do students use in their writing when reasoning about organic reaction mechanisms?
2. What changes do students make in the features present in their writing after peer review and revision?
3. How are students' revisions linked to the components of the peer review process?

7.6 Methods

7.6.1 Instructional setting

This research took place within a second-semester organic chemistry laboratory course at a large, Midwestern research university. The laboratory course included a weekly, one-hour lecture component that covered content and procedures relevant for the weekly, four-hour laboratory component. The lecture component was taught across three sections by faculty and postdoctoral instructors, while the laboratory component was taught across multiple smaller sections by graduate student instructors. The coursework included a laboratory notebook, quizzes, and three writing assignments. The writing assignments accounted for thirty percent of students' final grade for the course. Explicit instruction on using the EPF and interpreting RCDs took place near the beginning of the prerequisite first-semester organic chemistry lecture course (typically introduced during the second and fifth weeks of the course, respectively). RCDs are typically covered alongside substitution, elimination, and addition reactions. Instruction across the lecture and laboratory sequence incorporated using the EPF and RCDs to explain relevant phenomena; as such, students were expected to have enough familiarity with these representations to complete the WTL assignment described below. The course was affected by the onset of the COVID-19 pandemic, and the final weeks of the course were completed remotely. Therefore, the described WTL assignment and associated data collection took place entirely during remote instruction. All WTL assignments in the course were already administered asynchronously online through the learning management system, so no change was needed to the WTL implementation when changing to remote instruction.

7.6.2 Writing-to-learn assignment and implementation

The WTL assignment for this study was the final of three. The writing task was designed to afford students the opportunity to practice using the EPF and RCD representations to explain relevant phenomena. The assignment had the specific goal to support students in developing representational competence, particularly the abilities to (1) use representations to support their reasoning and (2) to select the appropriate representation for a task.[14] The assignment introduced students to a triazabicyclodecene (TBD) catalyzed intramolecular aldol reaction and two of its possible mechanistic pathways as identified by Hammar et al.[69] The EPF schemes and RCDs were provided for both mechanisms, as presented in Figure 7.1. The focus of analysis for this study was the portion of the assignment that required students to identify which of the two mechanistic pathways they thought to be the most likely and to explain their choice. As described in the conclusion to the article by Hammar et al., the most likely pathway is Mechanism A, based on the density functional theory (DFT) calculations used to determine the energy values represented in the RCDs.[69] Note that students were not expected to write about DFT in their responses but that students were expected to be able to provide a response based on the information given in the assignment. The assignment was designed to incorporate features demonstrated to support students' learning, including the opportunity to apply content knowledge to a meaning-making writing task and structures for peer interaction and revision.[45,46,54] The full text of the assignment is available in 7.11.1 Appendix 1.

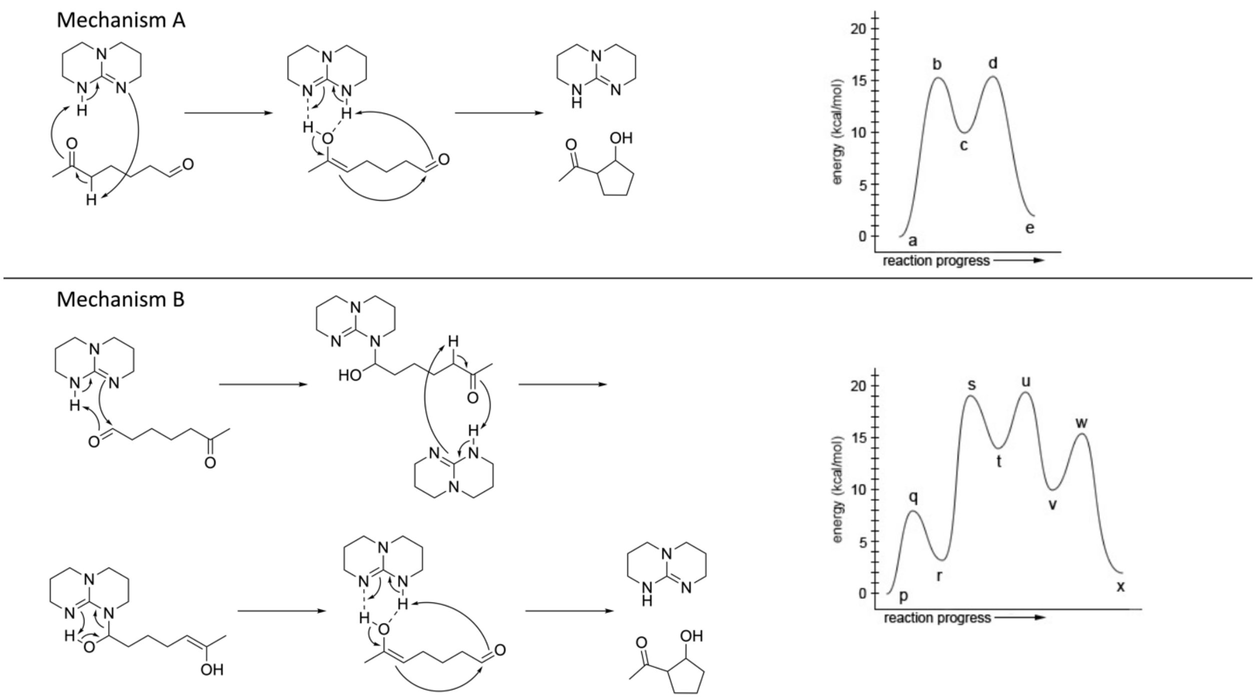


Figure 7.1 The EPF schemes and RCDs provided for both mechanisms within the WTL assignment, as shown to students.

The assignment was available to students in the online learning management system, and students had one week to submit an initial draft. Students then underwent an automated, double-blind peer review process in which each student provided and received feedback from typically three peers. Peer review was content-focused in that students provided feedback in alignment with the concepts targeted by the assignment. Students provided peer review comments by responding to the questions in the peer review guidelines (available in 7.11.1 Appendix 1). Following the peer review process, students had three days to revise their response and submit final drafts. Throughout the weeks the assignment was open, students had access to support from undergraduates who were previously successful in the course and trained as writing fellows to support students with the WTL assignments. The peer review process and availability of writing fellows were intended to provide structure for interactive writing processes and to encourage metacognition and reflection.[45,46] The assessment process for the assignment was independent of the presented analysis, though students were assessed based on whether they incorporated reasoning rather than the correctness of their response.

7.6.3 Participants and data collection

The participants for this study included 456 students out of the 771 students who received a final grade for the course. The 456 students were those who consented to participate in the study, submitted a first and revised draft, and both provided and received peer review feedback to/from other consenting students. The data collected for this study included the first drafts, peer review comments, and final drafts for the WTL assignment. Data were collected following Institutional Review Board approval for human subjects research.

7.6.4 Data analysis

Data was analyzed using a mixed methods approach, through which qualitative analysis of students' work was followed by quantitative transformation for further statistical analysis.[70] The mixed methods approach was chosen to enable us to view the data through multiple approaches and perspectives, allowing for better understanding of the complexities present in students' writing and their peer review interactions.[71] The qualitative and quantitative analysis stages are described below.

Qualitative analysis. The qualitative analysis took part in three stages: (1) identifying the mechanism students chose as the most likely in their initial and revised drafts, (2) qualitatively coding the initial and revised drafts for features of the mechanistic representations students wrote about to guide their choice, and (3) qualitatively coding peer review comments for whether students indicated agreement or disagreement with the mechanism chosen as the most likely.

The three coding schemes for the qualitative analysis are presented in 7.11.2 Appendix 2. All coding schemes were developed by two members of the research team (FW and GP). The first coding scheme was used to identify students' initial and revised responses for whether the student selected Mechanism A, Mechanism B, both, or neither as the most likely mechanism (7.11.2 Appendix 2, Table 7.4). At this stage of the analysis, responses from 456 students (both initial and final drafts for a total of 912 drafts) were analyzed. The second coding scheme was developed through inductively coding students' initial and revised drafts.[70] The inductive coding scheme was developed to identify how students were using features of the mechanistic representations (e.g., peaks on the RCDs or functional groups in the EPF schemes) within their justifications for the reaction pathway they indicated as the most likely to occur. Through multiple rounds of open coding and discussions with the research team, the coding scheme was revised and modified until saturation was reached and all codes were clearly defined (7.11.2 Appendix 2, Table 7.5). The coding scheme sought to identify the different aspects of students' writing that related to the features of mechanistic representations and/or their reasoning for their selection of the most likely pathway. For example, some codes captured students' use of specific words or phrases within their justification (e.g., the "thermodynamics" and "kinetics" codes), while others captured students' more specific reasoning for their selection of the most likely mechanistic pathway (e.g., the "general energy" and "counting" codes). Responses from 164 students (both initial and final drafts, for a total of 328 drafts and 1594 sentences) were analyzed with the finalized coding scheme. This coding was done on a sentence level using NVivo 12,[72] with the allowance that each sentence could be coded with all applicable codes. Two examples of a coded response (initial and revised drafts) are provided in 7.11.3 Appendix 3, Figure 7.9. The third coding scheme was developed to analyze the peer review comments students wrote in response to the fourth peer review criterion, which asked students to comment on whether the author selected the appropriate choice for the most likely mechanism. The coding scheme was used to categorize the comments as providing agreement, disagreement, or a neutral response (7.11.2 Appendix 2,

Table 7.6). The peer review comments received by all 456 students were analyzed (a total of 1361 comments).

Reliability. Efforts were taken to establish reliability throughout the qualitative analysis process.[73] For each stage, a subset of at least 20% of the total analyzed data was independently analyzed by two authors (FW and GP for stages 1 and 3; FW and MP for stage 2). The percent agreement and an appropriate IRR measure was calculated for each stage (Table 7.1). The calculated agreement measure for each stage of analysis indicates strong agreement.[73]

Table 7.1 Details for the efforts to establish reliability for each stage of the analysis

Data analysis stage	***N* (students)**	**Data analysed**	***N* (analysis units)**	**Reliability subset**	**Percent agreement (%)**	**IRR measure**
(1) Mechanism choice	456 students	Initial and revised drafts, at the document level	912 drafts	280 drafts (31%)	99	0.94[a]
(2) Features of mechanistic representations	164 students	Initial and revised drafts, at the sentence level	1594 sentences	334 sentences (21%)	82	0.80[b]
(3) Peer review comments	456 students	Peer review comments received	1361 comments	277 comments (20%)	90	0.80[a]

[a] Cohen's kappa, for when only one code is applied to each unit of analysis.[74] [b] Fuzzy kappa, for when more than one code can be applied to a single unit of analysis.[75]

Quantitative analysis. The results of the qualitative coding were transformed into quantitative data for further analysis.[70] The sentence-level coding from the second coding scheme of the qualitative analysis was used to determine frequencies with which each code appeared in each student's first and second drafts (since each code could be applied to multiple sentences and each sentence could have multiple codes). Students' sentence-level revisions were represented by subtracting the frequency with which each code appeared in their first drafts from the frequencies in their revised drafts. These data were used to perform the quantitative analyses focused on the features students included in their initial drafts and revisions, completed in R using RStudio.[76] Statistical significance for all analyses was set at $\alpha = 0.05$.

The statistical analyses for initial drafts and revisions were performed to identify sentence-level differences between groups of students based on the mechanism they selected as the most likely. To analyze initial drafts, we grouped students by whether they selected Mechanism A or Mechanism B. This allowed for identifying the connections students made between the features of representations (identified with the coding scheme in Table 7.5) and the mechanistic pathway they identified as most likely (identified with the coding scheme in Table 7.4). To analyze students' revisions, students were grouped by degree of global revisions, which are revision activities beyond the sentence level.[66] For the context of this study, global revisions were characterized by whether students revised to select a different mechanism as most likely (Table 7.2).

Table 7.2 Groupings of students by global revisions

Global revision group	Mechanism selected, initial draft	Mechanism selected, revised draft
1	A	A
2	A	B
3	B	A
4	B	B

The statistical analyses were selected and conducted as described by Sheskin.[77] Statistics first involved Shapiro–Wilk tests for normality, which indicated the distributions for the total number of codes in students' initial drafts were non-normally distributed ($W = 0.93$, $p < 0.001$). Hence, non-parametric tests were used for each analysis. Next, Mann–Whitney U tests were used to identify differences between the two groups of students on their initial drafts. Finally, Kruskal–

Wallis one-way analysis of variance by ranks tests were used to identify differences in sentence-level revisions between the four groups of students based on their global revisions. For statistically significant results on the Kruskal–Wallis tests, *post hoc* pairwise Mann–Whitney *U* tests were performed to identify the specific revision groups between which the differences were significant. For all tests in which multiple hypothesis tests were conducted simultaneously, *p*-values were corrected using Bonferroni's procedure to adjust for the family-wise Type I error rate.[77]

Lastly, we used logistic regression analysis to identify relationships between the global revisions and the two components of the peer review process (receiving feedback and reading other students' drafts). The logistic regression models took the general form of

Revisions = Initial Draft Mech. + Comments Received + Drafts Reviewed

where *Revisions* is the outcome variable capturing whether students made global revisions (i.e., whether students chose a different mechanism as most likely in their revised draft). This variable was coded as *Revisions* = 1 if students made global revisions (i.e., for groups 2 and 3 in Table 7.2) and as *Revisions* = 0 if students did not make global revisions (i.e., for groups 1 and 4 in Table 7.2). The *Initial Draft Mech.* predictor variable served as a binary indicator of the mechanism students selected as most likely in their initial draft. Note that students who selected both or neither mechanism in either draft were excluded from all regression models. The *Comments Received* predictor variable indicates the number of peer review comments students received that included a disagreement with the mechanism selected as most likely. Similarly, the *Drafts Reviewed* predictor variable indicates the number of drafts students reviewed that selected a different mechanism as most likely. Three logistic regression models were calculated and are described in detail in the results. Odds ratios were used to interpret the logistic regression models, which were calculated by exponentiating the coefficients of non-interaction terms. Odds ratios are interpreted as the factor by which each predictor variable influenced the odds of students making global revisions.[78]

7.7 Results

The central goal of this research is to identify how students use the two common representations of organic reaction mechanisms (RCDs and EPFs) to reason in writing about the likelihood of alternative reaction mechanisms. This analysis is set within students' responses to a WTL assignment in which they had the opportunity to revise their responses after undergoing a peer review and revision process. As such, this study also aims to identify how students' reasoning changes following peer review and revision. We first present the results of the initial stage of our analysis, in which we identified the mechanistic pathway students selected as most likely in their initial and final drafts. Following this, we address our three research questions to (1) describe the features of RCDs and EPFs students used to justify their decisions in their initial drafts, (2) describe the changes students made in their explanations after peer review and revision, and (3) identify the degree to which components of the peer review process—both receiving and providing feedback—influenced students' revisions. The results are interpreted and situated within the literature and theoretical frameworks in the discussion section.

7.7.1 The mechanistic pathway students selected as most likely in their initial and final drafts

As presented in Figure 7.2, most students selected Mechanism A as most likely in their first draft (n = 336), while most remaining students chose Mechanism B (n = 114). Few students did not clearly indicate a choice (n = 6). After the peer review process, few students who initially chose Mechanism A revised to choose Mechanism B in their revised draft (n = 5). In contrast, slightly less than half of the students who initially chose Mechanism B revised to choose Mechanism A in their revised draft (n = 46).

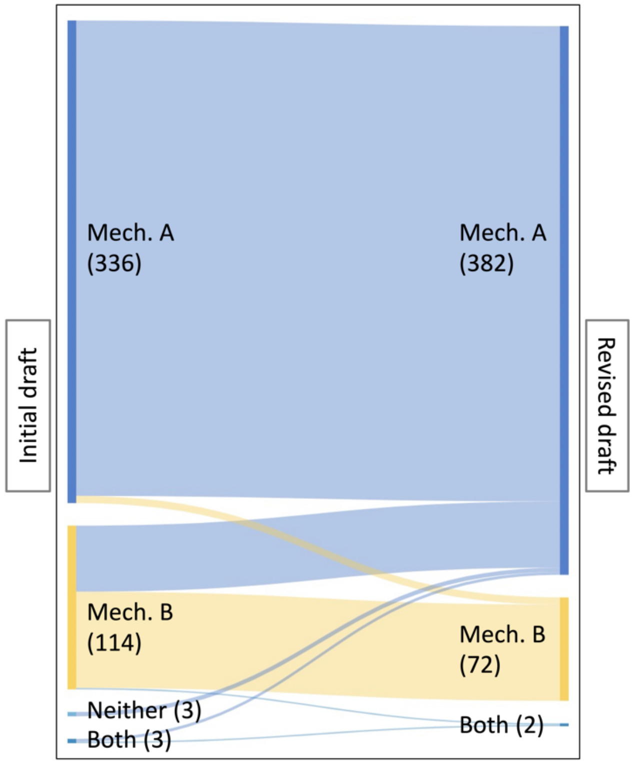


Figure 7.2 Sankey diagram representing the pathway students selected as most likely in their initial and revised drafts, illustrating the proportion of students making different types of global revisions. The total sample size reflected in the diagram is $N = 456$.

7.7.2 What features of the RCDs and EPFs do students use in their writing when reasoning about organic reaction mechanisms?

We present our analysis of the features present in students' initial drafts as captured by the qualitative coding process to address this research question. Each code corresponds to one or both of the mechanistic representations students used to support their identification of the most likely mechanistic pathway. We sought to identify differences between students who identified different mechanisms as most likely. First, there is no significant difference between the total number of codes between students selecting the different mechanisms (Mann–Whitney *U* test, $W = 2331$, $p = 0.25$). This finding indicates that students selecting one mechanism did not tend to incorporate more features in their writing, as captured by the coding scheme, compared to students selecting the other mechanism. Next, we sought to identify if there were differences in the specific features students incorporated based on the chosen mechanism. The average frequency of each code appearing across students' initial drafts is presented in Figure 7.3. The significance levels are shown from the outcome of the Mann–Whitney *U* tests for differences in students' first drafts depending on whether they selected Mechanism A or Mechanism B as the likely mechanistic pathway. The relevant data for this research question, including mean values, standard deviations, and Bonferroni corrected *p*-values, is also presented in 7.11.4 Appendix 4, Table 7.7. These results indicate the specific features of students' writing they incorporated to support their choice of the likely mechanistic pathway and which features were significantly different among students selecting the different mechanisms as most likely (with the significant codes being *TBD adding* and *functional group*).

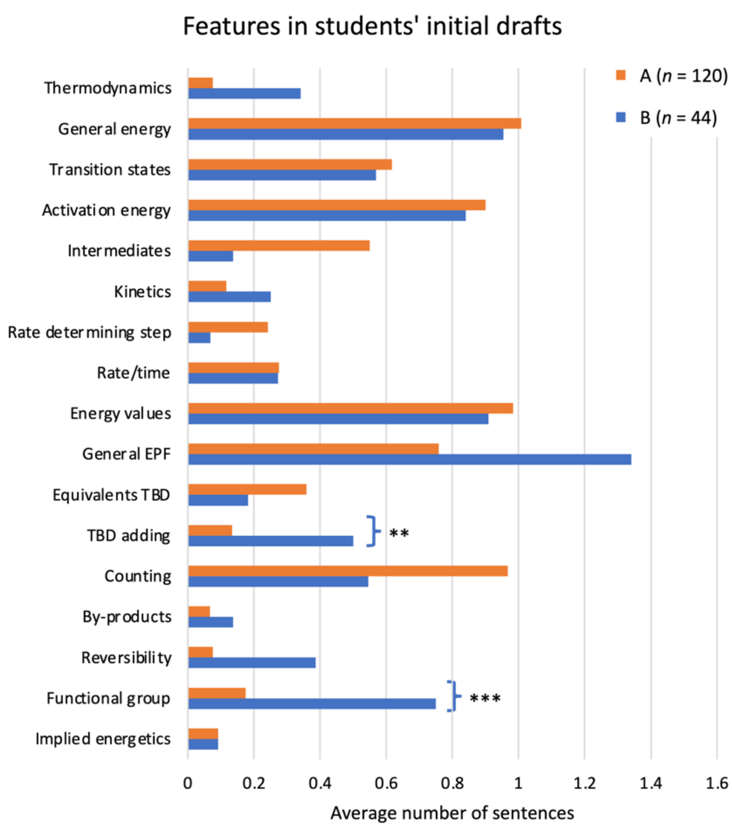


Figure 7.3 The average frequency of sentences for each code appearing in students' initial drafts, separated by whether students indicated Mechanism A or Mechanism B as most likely. Definitions for each code can be found in 7.11.2 Appendix 2, Table 7.5. Significant differences between groups are indicated with $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

7.7.3 What changes do students make in the features present in their writing after peer review and revision?

We performed a similar analysis to investigate the features students used to guide their choice of the most likely mechanism in their revised drafts. For each student, the frequency of each code applied to their initial draft was subtracted from the frequency of each code applied to their revised draft, resulting in a value that indicates the frequency with which each code was added (or removed) upon revision. The average change in the frequency of each code appearing in students' revisions, across the four revision groups in Table 7.2, are presented in Figure 7.4. The significance levels are shown from the outcome of the Kruskal–Wallis tests for differences in students' revisions across the four groups. The relevant data for this research question, including mean values, standard deviations, and Bonferroni corrected *p*-values, is also presented in 7.11.4 Appendix 4, Table 7.8. For the three codes with significant differences across revision groups, the results of the *post hoc* pairwise Mann–Whitney *U* tests are presented in 7.11.4 Appendix 4, Table 7.9. These provide further specification about the groups between which the differences in revisions were significant. Altogether, the results for this research question indicate the specific codes with significant differences in the frequencies of revisions across the four revision groups (with the significant codes being *TBD adding*, *counting*, and *functional group*).

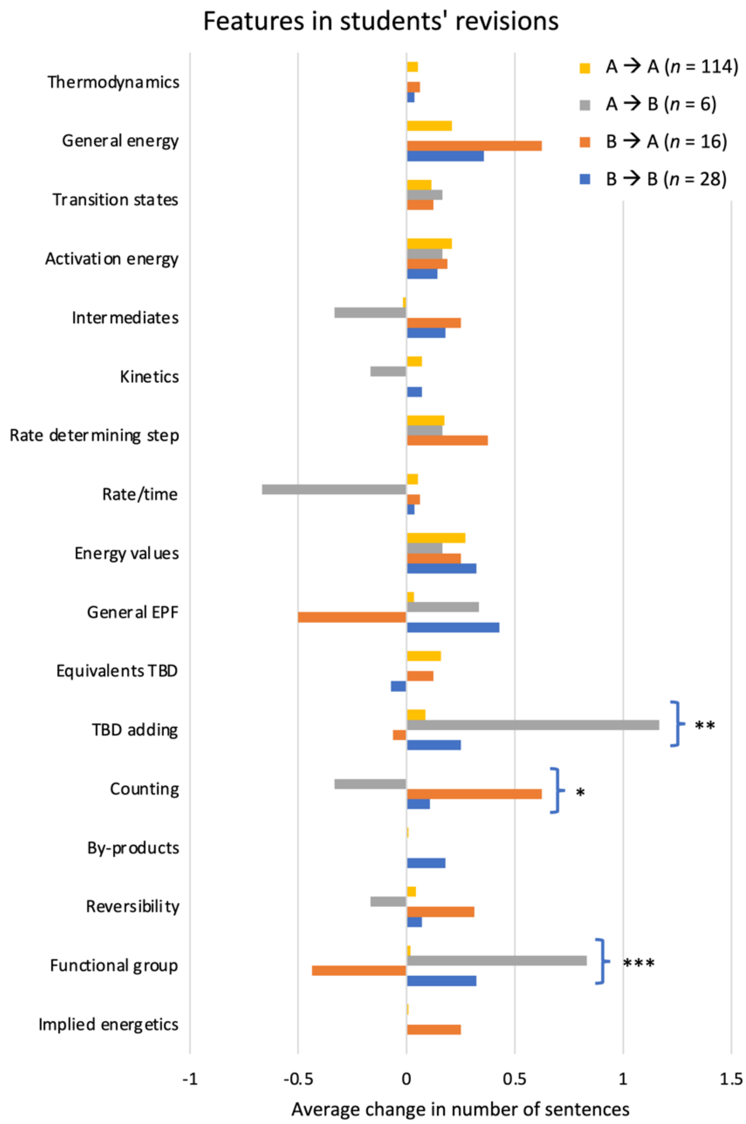


Figure 7.4 The average change in frequency of sentences for each code appearing in students' revisions, separated by the nature of students' global revisions. Definitions for each code can be found in 7.11.2 Appendix 2, Table 7.5. Significant differences between groups are indicated with $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

7.7.4 How are students' revisions linked to the components of the peer review process (both receiving and providing feedback)?

The final stage of the analysis sought to identify if and how the global revisions students made were connected to the two components of the peer review process: receiving feedback from peers and reviewing peers' drafts. To do this, we performed three logistic regression analyses. The outcome variable for all regressions was whether students revised to select another mechanism. The predictor variables for the regressions were (1) the mechanism students indicated as most likely in their initial drafts, (2) the instances of disagreements encountered in peer review comments received, and (3) the instances of reviewing drafts that selected the opposite mechanism as most likely. The results of the three logistic regressions are presented in Table 7.3. The descriptive statistics for the variables included in the regression analysis are in 7.11.4 Appendix 4, Table 7.10.

Table 7.3 Summary of the logistic regression models

Predictor	**Coeff. (st. err.)**	**Exponent of coeff.**	***p*-value**
Model 1. Logistic regression analysis without interaction terms			
Initial draft mech. (A = 0, B = 1)	1.44 (0.58)	4.24	0.013*
Comments received (freq. of disagreement)	0.88 (0.24)	2.41	<0.001***
Drafts reviewed (freq. of disagreement)	1.45 (0.30)	4.26	<0.001***
Intercept	–6.58 (0.81)	0.00	< 0.001***
Model 2. Interaction analysis with students who initially selected Mechanism A as most likely as the reference group			
Initial draft mech. (A = 0, B = 1)	2.14 (1.71)	8.51	0.210
Comments received (freq. of disagreement)	1.16 (0.72)	3.19	0.108
Drafts reviewed (freq. of disagreement)	1.55 (0.50)	4.72	0.002**
Initial draft mech. × Comments received	–0.33 (0.76)	0.72	0.665
Initial draft mech. × Drafts reviewed	–0.19 (0.63)	0.83	0.766
Intercept	–6.99 (1.29)	0.00	< 0.001***
Model 3. Interaction analysis with students who initially selected Mechanism B as most likely as the reference group			
Initial draft mech. (A = 1, B = 0)	–2.14 (1.71)	0.12	0.210
Comments received (freq. of disagreement)	0.82 (0.25)	2.29	< 0.001***
Drafts reviewed (freq. of disagreement)	1.37 (0.38)	3.92	< 0.001***
Initial draft mech. × Comments received	0.33 (0.76)	1.39	0.665
Initial draft mech. × Drafts reviewed	0.19 (0.63)	1.21	0.766
Intercept	–4.85 (1.13)	0.01	< 0.001***
N = 449. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$			

Model 1 included only the three predictor variables, with *Initial Draft Mech.* = 0 for students who indicated Mechanism A and *Initial Draft Mech.* = 1 for students who indicated Mechanism B as most likely. Model 1 indicates that all three predictor variables significantly predict whether students will make global revisions to select the opposite mechanism in their final draft. The exponent of the coefficients (the odds ratios) of the predictor variables for Model 1 indicate the size of the influence. For example, the odds ratio of 4.24 for initial draft mechanism choice in Model 1 indicates that a student who initially selected Mechanism B as most likely has a 4.24 odds of making global revisions over a student who initially selected Mechanism A. Similarly, the odds ratio of 2.41 for the *Comments Received* variable in Model 1 indicates that a one unit increase in the number of peer review comments that state a disagreement increases the odds that the student will make global revisions by a factor of 2.41.

Model 2 and Model 3 included interaction terms to investigate the relationship between *Initial Draft Mech.* with the *Comments Received* and *Drafts Reviewed* variables. Model 2 used students who initially selected Mechanism A as the reference group (i.e., *Initial Draft Mech.* = 0 for students who initially indicated Mechanism A as most likely). Because of this, the exponent of the coefficients for the non-interaction terms (the odds ratios) are interpreted conditionally for students who initially selected Mechanism A.[78] For example, the odds ratio of 4.72 for the *Drafts Reviewed* variable in Model 2 indicates that, for students who initially selected Mechanism A as most likely, each one unit increase in the drafts reviewed that selected Mechanism B increases the odds of global revisions by a factor of 4.72.

For Model 3, the odds ratios are interpreted conditionally for students who initially selected Mechanism B, as the model was calculated with students who initially selected Mechanism B as the reference group (i.e., *Initial Draft Mech.* = 0 for students who initially indicated Mechanism B as most likely). For example, the odds ratio of 3.92 for the *Drafts Reviewed* variable in Model 3 indicates that, for students who initially selected Mechanism B as most likely, each one unit increase in the drafts reviewed that selected Mechanism A corresponds to an increase in the odds of global revisions by a factor of 3.92. Note that the exponents of the interaction terms for Models 2 and 3 are not odds ratios and are therefore not to be interpreted in the same manner.[78]

Model 2 results indicate that the only significant predictor variable for students who initially selected Mechanism A is *Drafts Reviewed*. However, the results for Model 3 indicate that both the *Comments Received* and *Drafts Reviewed* variables are significant predictors of global revisions for students who initially selected Mechanism B. Together, the results of all three models indicate a positive direction of influence for both components of the peer review process on the outcome of revising to select a different mechanism as most likely. That is, encountering a disagreement in both comments received or drafts reviewed influenced students' revisions to select the mechanism that matched their peers. Thus, the results of Model 2 and Model 3 specifically provide insight into the nature of the interaction between the mechanism students initially selected as most likely and the influence of both components of the peer review process.

7.8 Discussion

The presented results indicate the general trends across students' initial responses, revisions, and interactions in the peer review process. These trends point to the key findings and claims we can make from our analysis, considering the existing literature and theoretical frameworks guiding this study. The following discussion is organized by the key findings and claims.

7.8.1 Students largely selected the favored mechanistic pathway as the most likely mechanism in both their initial and revised responses

This finding is promising as, at the end of the WTL process, most students successfully selected the most likely mechanistic pathway (n = 382, Figure 7.2). Furthermore, most students who revised their choice of the most likely mechanistic pathway transitioned to select the appropriate pathway (n = 51 of 57 students, Figure 7.2). That students made global revisions provides evidence that the WTL process encouraged reflection and revision for these students, an intended goal of WTL assignments.[45,46,66] This result provides further evidence for the value of WTL in organic chemistry,[19,50,60,61] while extending the findings of prior studies by demonstrating students' engagement with WTL on a task that required consideration of two reaction mechanisms, represented by both the EPF and RCDs. Additionally, this finding provides evidence for using the WTL process with peer review and revision to support students' conceptual engagement within the organic chemistry course context.

7.8.2 Students across the dataset incorporated features from both representations in their responses

The closer analysis of students' initial and revised drafts further supports the claim that the WTL assignment supported students' conceptual engagement. The results for research questions two and three indicate that students across the dataset incorporated evidence from both the RCDs and EPF schemes to support their claim of the most likely mechanistic pathway. This indicates that the WTL assignment supported students' use of representations in their reasoning, moving beyond engaging students with representations as simply being *of* a phenomenon.[9,14] However, the findings indicate differences among students both within and between the groups who selected different mechanisms as most likely, suggesting nuance in students' developing representational competence. Notably, different students who selected the same mechanism (e.g., different students who selected Mechanism A at some stage of the WTL process) occasionally exhibited different reasoning. This finding is described in detail in the following paragraphs and provides support for the calls in the literature to emphasize and evaluate the process of students' reasoning rather than the final outcome or product of their reasoning.[19,38,40,79]

7.8.3 Students who selected Mechanism A as most likely reasoned by appealing to both chemically accurate and chemically inaccurate reasoning

In general, students who selected Mechanism A as most likely included two specific reasons: that Mechanism A had lower activation energy and that Mechanism A had fewer steps. For instance, one student reasoning with activation energy wrote: "Mechanism A is more likely to occur because it has a lower activation energy and proceeds through lower energy intermediates and transition states than Mechanism B." Students who used this reasoning demonstrated an appropriate interpretation of the RCD representation. However, this finding is complicated because students often discussed energy in broader terms (as captured by the *general energy* code). Students' explanations that referred to energy both in general terms and specifically with the phrase "activation energy" align with the research from both Lamichhane et al. and Popova and Bretz by suggesting that some students may not have an understanding of activation energy in alignment with chemists' interpretations.[2,4] The result of the present study extends these findings by illustrating that some students could identify the appropriate concept and representation for completing the writing task (i.e., energetics derived from the RCDs) but may not have been able to connect the concept to the specific, appropriate feature of the RCD representation (i.e., peaks representing the activation energy). Students' continued discussion of energy in general terms, rather than specific terms, after the cognitive processes of writing and revising indicates an area where instruction can be improved to support students' representational competence.[14,64–66]

The other common reasoning among students who selected Mechanism A was based on the number of steps in the reaction. For example, one student wrote: "This mechanism also has fewer steps which makes it more favorable when synthesizing for real application." This finding suggests a different approach to the task in comparison to the reasoning based on energetic considerations. Specifically, students who reasoned based on the number of steps demonstrated focus on a representational surface feature that could be drawn from either the RCD or EPF.

Students' writing that appealed to the surface features of the representations aligns with prior studies indicating students' reliance on surface features for both RCDs[2–5,31,32] and the EPF.[16–19,23,26,28] Students' reliance on representational surface features is valuable to identify, as a key component of developing representational competence is being able to interpret the chemical meaning of representations when supporting their reasoning.[14] Specifically, the reasoning that the likelihood of alternative reaction mechanisms is based on the number of mechanistic steps is notable, as it represents a naïve view of mechanism that does not incorporate chemical reasoning. This approach to reasoning is in alignment with the type of rule-based reasoning that neglects chemical understanding, which has been reported in the literature.[24,33,34,38,40] Furthermore, that the WTL assignment elicited this type rule-based of reasoning suggests the value of WTL for eliciting students' understanding that may be difficult to elicit through other assessment approaches, which has been suggested in a prior study of WTL.[48]

7.8.4 Students who selected Mechanism B as most likely reasoned by appealing more to the EPF than students who selected Mechanism A as most likely

Students who reasoned that Mechanism B was most likely incorporated reasoning more focused on the EPF, particularly with respect to the first step of Mechanism B (in which the TBD catalyst acts as a nucleophile and adds to the aldehyde). The only two significantly different codes identified in research question one related to this reasoning. This result suggests that students who selected Mechanism B were also writing about the features that guided other students' choice of Mechanism A as most likely (e.g., features as captured by the codes *general energy*, *activation energy*, or *counting*). However, the students who selected Mechanism B also incorporated significantly more reasoning captured by the codes *TBD adding* and *Functional group*. The *TBD adding* code is exemplified by one student's response that.

> "Mechanism B is the most likely pathway because unlike Mechanism A, TBD reacts with the aldehyde side of the ketoaldehyde first which prevents the possibility of the aldehyde reacting with TBD and forming an enol."

Similarly, another student wrote:

> "In Mechanism B, however, the first step eliminates this type of reactivity at the aldehyde since of the binding TBD which makes it an alcohol. This is much less likely to occur at

the ketone due to steric hindrance, therefore making Mechanism B more selective than Mechanism A."

This reasoning is drawn from the EPF and reflects the idea that TBD must first react with the aldehyde to act as a protecting group before the reaction can occur at the more sterically hindered ketone. Such reasoning focused on sterics has been demonstrated by students in a prior study.[27] While this reasoning about the EPF is reasonable—and can be supported by comparing the transition state energies for the initial steps for the two reaction mechanisms—it is notable that students based their selection of the most likely pathway on this feature of the EPF rather than the more appropriate features of the RCDs. This reflects students' appropriate interpretation of the EPF representation, but their inappropriate selection of the representation most suited for the task of selecting the most likely reaction pathway. Hence, these students exhibited some aspects of representational competence, but not the ability to select the appropriate representation of a phenomenon for a specific task.[14]

The other code significantly more common among students who selected Mechanism B as most likely was the *Functional group* code. This is exemplified by a student who wrote: "I believe that Mechanism B is more likely to occur because aldehydes are usually more reactive than ketones." This reasoning, similar to the *TBD adding* code, is focused on the idea that aldehydes are more reactive than ketones—but without the explicit reference to the steric hinderance argument. These students suggested that because aldehydes are more reactive, the reaction that starts with the aldehyde is more likely. This reasoning aligns with prior studies of students' reasoning with functional group reactivity trends rather than the actual functions of said functional groups.[23,40] Furthermore, this reasoning focused primarily on reactivity trends is aligned with the rule-based reasoning strategy demonstrated by students in prior studies of their reasoning in organic chemistry.[24,33,34,38,40] In addition, the focus on functional group is explicitly tied to features of the EPF (i.e., identifying functional groups) and plausibly tied to students' connection-making between the EPF and RCDs (i.e., examining the reacting functional group on the EPF and the corresponding transition state energy on the RCD). Nevertheless, students exhibiting this reasoning included additional focus on the EPF, the representation less suited for the task of selecting the more likely reaction pathway—suggesting that these students are still developing their ability to select the most appropriate representation for a given task.[14]

7.8.5 Students' revisions revealed similar trends in reasoning for selecting both Mechanism A and Mechanism B, while both reducing and eliciting students' inaccurate reasoning

Similar trends as those identified from the sentence-level analysis of students' first drafts were identified in students' revisions. Students across the dataset generally revised to incorporate more features in their responses, with some exceptions (as seen in Figure 7.4). However, the differences between students adding or removing features were not significant for most features identified in students' writing. Three features, however, were significantly different across the different groups of students based on their global revisions: the features captured by the *TBD adding*, *Functional group*, and *Counting* codes. These features mirror the trends identified in students' initial drafts. Students who initially selected Mechanism B and revised to select Mechanism A tended to revise their writing to remove the features corresponding to the relative reactivities of aldehydes and ketones while adding the reasoning based on counting the number of steps. It is notable that, through the cognitive processes of writing,[64–66] many students reduced the prevalence of their reasoning about the EPF schemes that was less appropriate for the writing task when revising to select Mechanism A. However, the reverse was true for students who initially selected Mechanism A and revised to select Mechanism B. This trend follows from the results of analyzing students' initial drafts by suggesting that these features guided students' decisions to identify Mechanism B as more likely. Altogether, the nature of students' revisions suggests how the peer review and revision process might serve to elicit some students' inappropriate reasoning that was not elicited in their initial drafts. This finding extends the WTL literature to support the notion that peer review and revision are useful for identifying reasoning that might be challenging to elicit through other means of assessment.[48] Furthermore, this finding also indicates that while students may select the more likely mechanistic pathway, they do not always exhibit accurate chemical reasoning even after the complete WTL process. This supports the emphasis present in the literature to focus on eliciting, emphasizing, and evaluating students' reasoning itself rather than the products or outcomes of their reasoning.[19,38,40,79]

7.8.6 Students' global revisions are influenced by both reviewing their peers' work and by receiving peer review comments

The final set of findings relates to students' interactions within the peer review process and how it influenced students' decisions to make global revisions. The logistic regression models indicate similar trends and key differences based on the mechanistic pathway students indicated as most likely in their initial drafts. For both groups of students, disagreements in the drafts reviewed significantly predicted students' decisions to switch which mechanism they selected as most likely within their revisions. Furthermore, the odds ratios were higher for the disagreements in drafts reviewed compared to disagreements in comments received. Together, each of these findings indicate that students are influenced to a higher degree by the drafts they review than by the comments they received. The finding that student are influenced more by the drafts they review was suggested in prior studies of peer review, both in STEM WTL contexts[53] and traditional writing courses.[56–58] However, for students who incorrectly selected Mechanism B as the more likely mechanism in their initial drafts, disagreements in the peer reviewer comments received were also statistically significant predictors of students' global revisions. The same was not true for students who initially selected Mechanism A. This result suggests two claims: (1) that peer review comments can influence students' decisions to revise (though a smaller effect than the drafts students review) and (2) that disagreements in peer review comments have more influence for students who initially display inaccurate reasoning. Altogether, the peer review analysis results indicate that students are influenced to make global revisions from both receiving comments and reviewing others' drafts, and that the influence differs based on which mechanism they selected as most likely in their initial response. This aligns with previous studies of writing assignments which emphasize how the social components of writing can influence students' engagement with all aspects of the WTL process, creating space to encourage reflection and metacognition.[45,46] Furthermore, this finding extends the literature by identifying a possible interaction between peer feedback received and the accuracy of students' reasoning in their initial drafts.

7.9 Limitations

There are limitations associated with this study. First, the study took place within one laboratory course at a single institution, and thus findings may not be generalizable to all students or instructional settings. Students' responses may have been influenced by the writing assignment taking place within a laboratory course; for example, writing assignments (e.g., laboratory reports) are more typical in laboratory courses compared to lecture courses at the study institution, so students may have been more receptive to completing the writing assignment in the laboratory course setting. Furthermore, data collection took place during remote learning due to the COVID-

19 pandemic, which may have influenced the results. For instance, students may have engaged in fewer informal discussions surrounding the WTL assignment (e.g., while waiting for class to start) than might have been expected with in-person learning. This may have influenced students' responses and/or the degree to which students engaged in the peer review process. There are additional limitations imposed by the approach to data collection and analysis. First, only students' responses to the WTL process were collected. These responses serve as evidence of students' reasoning, but no other data such as interviews were collected to triangulate the findings. Nevertheless, students' abilities to use representational features to make claims suggest evidence of students' reasoning through the cognitive processes of writing. Additionally, due to qualitatively coding on the sentence level where multiple codes could be applied to each sentence, the analysis is not suited for making claims about the different ways students reasoned within groups. For example, the analysis does not allow for strict categorization of students between the different reasonings for selecting Mechanism A (i.e., based on activation energy *versus* based on counting the number of steps). However, this method of analysis was employed to account for all features students incorporated in their writing and to gain detailed perspective on the nature of students' revisions. Furthermore, there are limitations associated with quantifying and performing statistical tests on qualitative data. Specifically, any biases that may have affected the qualitative data analysis would be carried through into the quantified data and ensuing statistical analyses, which may have influenced the nature of the results. A final limitation of the analysis is that the goal was to generalize students' responses across the hundreds of participants in the study. Because of this, any individual student's response might represent reasoning that differs from the general trends identified in the analysis.

7.10 Conclusions and implications

This study describes our analysis of all components of a WTL assignment (students' initial drafts, peer review, and revised drafts) in which students were given a writing task to consider two common mechanistic representations in organic chemistry (the EPF and RCDs) in their reasoning about the likelihood of two alternative mechanisms for a single transformation. The analysis indicated that many students correctly selected the mechanistic pathway accepted as most likely in both their initial and revised responses, with slightly less than half of the students who selected the less likely pathway revising to select the more appropriate choice. Students across the dataset drew on features of both the RCD and EFP representations in their responses. Students who selected the more appropriate pathway (Mechanism A) reasoned appropriately by appealing to the RCDs and comparisons of activation energy; students in this group also reasoned less appropriately about the number of steps in the mechanism. Students who selected the less likely pathway (Mechanism B) reasoned by appealing to the EPF, discussing both steric considerations and reflecting knowledge about general reactivity trends for aldehydes and ketones. These findings suggest students who incorporated more reasoning with the EPF tended to select the less likely pathway, indicating the need to develop their representational competence skill of selecting the most appropriate representation for the task of identifying the more likely reaction mechanism. Students' revisions revealed similar trends in reasoning, with some students incorporating revisions that reflected inappropriate reasoning that was not revealed in their initial drafts. Finally, the peer review analysis indicates the potential influences of peer review on students' revisions, providing evidence that reading drafts with different perspectives has a higher odds of influencing students' global revisions compared to receiving feedback with disagreements. Altogether, these findings extend the literature by providing insight into organic chemistry students' representational competence as presented in their responses to a WTL assignment. Further, the findings provide key implications for research and practice both for WTL interventions and for teaching organic chemistry.

7.10.1 Implications for research

The results of this study further the understanding of how students engage with representations in organic chemistry to support their reasoning about reaction mechanisms. The study expands the growing literature that utilizes writing analysis to access students' engagement with disciplinary skills.[19,43,44] The present study specifically used writing analysis to provide insight into students' representational competence through examining all components of a WTL assignment, a methodology that can similarly be used for future studies. Writing analysis of this type enables researchers to analyze the reasoning of large numbers of students participating in a course and is a strategy that can overcome some of the limitations of interview analyses which often include a smaller set of self-selected participants. Future research is merited to further investigate organic chemistry students' representational competence. While many existing studies provide insight into how students interpret the features of organic chemistry representations, the present study indicates how students use representations for a specific task. Future research is necessary to explore how students use representations for other tasks similar to the work of practicing chemists. Additionally, further research is necessary for investigating other aspects of students' developing representational competence, as outlined by Kozma and Russell,[14] such as how students make connections between representations in their reasoning. There is also a need for further research into the components of the WTL process, including studies that investigate more specifically what influences students to make both global and sentence-level revisions across the peer review process.

7.10.2 Implications for practice

There are a variety of implications for teaching associated with this study. First, the study provides details on students' representational competence with the primary representations for organic reaction mechanisms. Understanding how students think about these representations at the introductory level is important for knowing how students might think about and approach different problems when learning organic chemistry. For example, knowing that some students may think that reaction mechanisms with fewer steps are more likely is valuable for teaching other reaction mechanisms in organic chemistry where alternative mechanistic pathways have different numbers of steps, such as substitution and elimination reactions. This study also provides a WTL assignment that engaged students with developing their representational competence, specifically with the ability to select the appropriate representation when completing a specific task. Activities which engage students in this type of task are especially important, as this is a component of representational competence that may not often be emphasized in introductory organic chemistry.[15] The results of the analysis of students' responses to this assignment suggest that teaching should specifically target the different uses for different representations, in alignment with the representational competence framework and calls to teach the epistemic practices surrounding developing and using models.[6–9,14] Specifically, instruction could be improved by exposing students to both common representations of organic reaction mechanisms throughout the introductory organic chemistry curriculum with emphasis on what information each representation provides and how each representation can be used for different problems in organic chemistry. This study also provides insight into the value of using WTL assignments within the organic chemistry classroom, particularly in that writing assignments can elicit students' inappropriate or non-chemical reasoning that might be difficult to elicit through other assignments or assessments. This study specifically suggests the value of implementing peer review and revision with WTL assignments for identifying how students reason with multiple representations. Lastly, this study provides support for the key implication of evaluating students' reasoning itself rather than the product of students' reasoning, as the findings indicate that students can use inaccurate reasoning to arrive at correct answers.

7.11 Appendices

7.11.1 Appendix 1. WTL assignment and peer review criteria

Exploring possible reaction pathways for a catalyzed intramolecular aldol reaction

Ivermectin is a drug used to treat onchocerciasis, a parasitic disease commonly known as river blindness. While the disease is rare in the United States, it is especially prevalent in Ghana, where more than 15% of the population is affected. As a lab technician for Médecins Sans Frontieres (Doctors Without Borders), you have traveled to Ghana to collaborate on a study initiated by biochemists at the University of Ghana who are working to develop a more efficient synthesis of ivermectin. The biochemists you are working with have identified a new strategy to perform intramolecular aldol reactions that uses the catalyst triazabicyclodecene (TBD). The TBD-catalyzed aldol reaction could be used in the place of the traditional aldol reaction for an early synthetic step in the synthesis of ivermectin. Using TBD will replace the need of strong acids and bases in this synthetic step, which will limit undesired side reactions. An example of a TBD-catalyzed aldol reaction with a simplified starting material is shown in Figure 7.5.

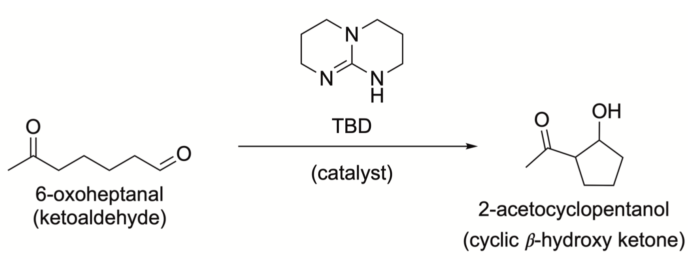


Figure 7.5 The intramolecular, TBD-catalyzed aldol reaction of 6-oxoheptanal produces 2-acetocylcopentanol.

The biochemists you are working for have asked you to research the mechanisms for the reaction. This will help them determine the feasibility of applying it to the synthesis of ivermectin.

You have identified two potential mechanistic pathways, shown below in Figure 7.6 and Figure 7.7.

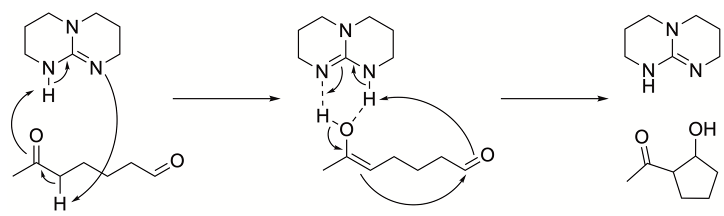


Figure 7.6 Proposed Mechanism A.

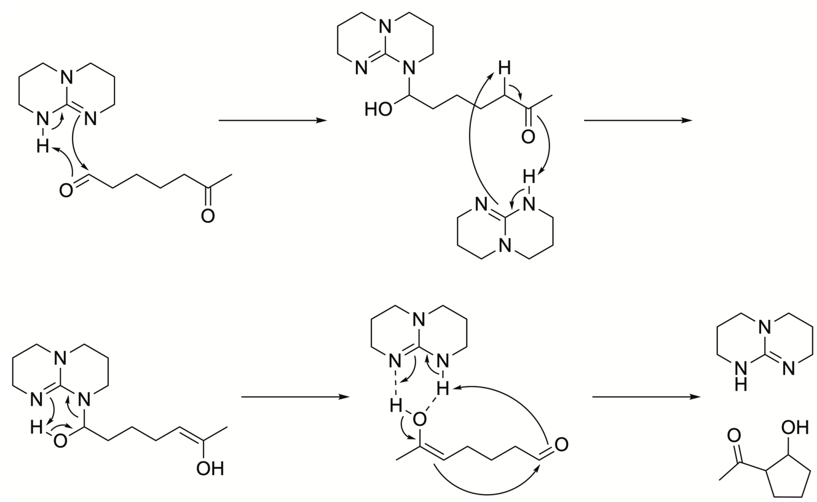


Figure 7.7 Proposed Mechanism B.

For each proposed pathway, you have performed computer simulations to determine their energy profiles. The results of your calculations are shown in Figure 7.8, where each reaction coordinate diagram is presented side-by-side.

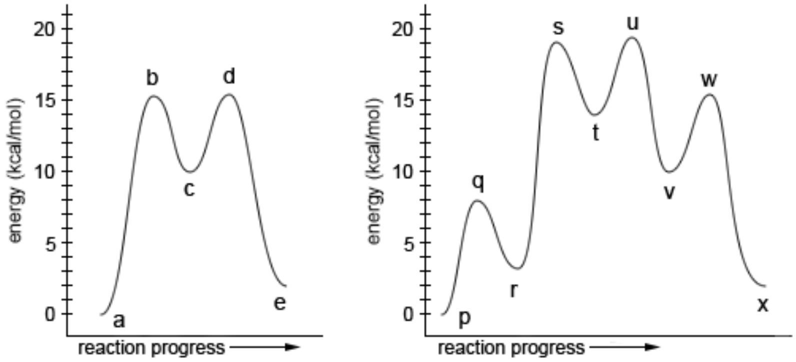


Figure 7.8 Reaction coordinate diagrams for Mechanism A (left) and Mechanism B (right). Note that claims about reaction times between Mechanism A and B can't be made since the units on the horizontal axes aren't specified.

At the end of the summer, you will write a brief report to summarize your findings, suggest the most likely pathway, and share your part of the project with the rest of the team. You should provide a detailed explanation of the mechanisms for both reaction pathways. Also, your argument for the most likely pathway should be supported by the mechanisms and the reaction coordinate diagrams. The report is directed toward the biochemists and other concerned parties who will use your recommendations to decide the feasibility of applying this reaction to the more complicated synthesis of ivermectin. Therefore, they may not be experts when it concerns mechanisms or organic-specific terms. Use clear and concise language, striking a balance between organic jargon and oversimplified explanations.

Your report should be approximately between 500–700 words (1–2 pages) in length. It should address the following points:

1. Discuss how each mechanism correlates with the corresponding energy diagram.
    a. Summarize the findings.
    b. Specifically, explain how the transition states and intermediates of the mechanisms correspond to features on the diagrams.
    c. Take care to translate which specific step in the mechanism corresponds to which specific feature of the associated reaction coordinate diagram.
2. Identify which reaction pathway you think is most likely to occur. You will be evaluated on the explanation of your choice, *not the choice itself*.

3. When discussing mechanisms, be sure to write about the structural features and electronics of the molecules involved. Include descriptions of how the molecules interact in the mechanism and how they change in structure as a result of their interactions.

You can and should include figures of schemes, structures, mechanisms, or reaction coordinate diagrams, if that supports your response. We suggest that you have the figure(s) in front of you—ready to color-code or mark-up in various ways—and that you use your visible thinking to guide your audience through your explanation. Any images that you include in your response, *including the figures in this prompt or those that you draw in ChemDraw or on paper*, must have the original source cited using either ACS or APA format. Given your audience, your written response should suffice so that the explanations can be understood without the figures. You will be graded only on your written response.

Peer review guidelines

- Print and read over your peer's essay to quickly get an overview of the piece.
- Read the essay more slowly keeping the rubric in mind.
- Highlight the pieces of texts that let you directly address the rubric prompts in your online responses.
- In your online responses, focus on larger issues (higher order concerns) of content and argument rather than lower order concerns like grammar and spelling.
- Be very specific in your responses, referring to your peer's actual language, mentioning terms and concepts that are either present or missing, and following the directions in the rubric.
- Use respectful language whether you are suggesting improvements to or praising your peer.

1. In what ways does the author discuss the structural features of the molecules and the changes that result from the interactions in Mechanism A? Suggest ways the author could improve their mechanistic description.
2. In what ways does the author discuss the structural features of the molecules and the changes that result from the interactions in Mechanism B? Suggest ways the author could improve their mechanistic description.

3. How does the author relate the mechanistic details to the corresponding energy diagram for each mechanism? Suggest specific ways the author could relate each mechanism to features of its energy diagram.
4. Which mechanistic pathway did the author choose as the most likely? State what choice you think the author made and whether or not you think the author made the correct choice. Provide an explanation for why you think this way.
5. How did the author justify their choice of the most likely mechanistic pathway? Suggest ways the author could use details from their mechanism and energy diagram descriptions to better explain their choice.

7.11.2 Appendix 2. Coding schemes

Table 7.4 Coding scheme for the first analysis stage, in which initial and revised drafts were coded for the mechanism students indicated as most likely

Code	Definition	Exemplar
Mechanism A	The student indicated Mechanism A as the most likely mechanism for the reaction.	"My belief is that the reaction undergoes mechanism A because it requires much less energy than mechanism B and does not form as many stabilized intermediates as mechanism B."
Mechanism B	The student indicated Mechanism B as the most likely mechanism for the reaction.	"Although it involves more steps, I believe that Proposed Mechanism B is more likely to occur."
Both mechanisms	The student indicated, in different parts of their response, both Mechanism A and Mechanism B as the most likely mechanism for the reaction.	"Based on the energy diagrams, it seems most likely that mechanism B occurs over mechanism A… As a result, mechanism A requires marginally less energy to progress past the rate-determining step and is more likely to occur than mechanism B."
Neither mechanism	The student did not clearly indicate either mechanism as the most likely mechanism for the reaction in any part of their response.	N/A

Table 7.5 Coding scheme for the second analysis stage, in which initial and revised drafts were coded at the sentence level for features guiding students' responses

Code	Definition	Exemplar
Thermodynamics	The student uses the word "thermodynamics" to describe the thermodynamics of the reactions.	"Thermodynamically speaking, both reactions are equally favorable since both have a $\Delta G°$ of 2 kcal/mol."
General energy	The student uses the word "energy" OR the student gives a generic description of the energy required for the reaction. Can include comparing "energy" between different points on the RCD.	"Mechanism A will be favored because it is much more energetically favored than Mechanism B."

Transition states	The student considers the highest energy or the transition state peaks on the RCDs, including phrases like "first peak."	"In determining the success of either mechanism in producing the desired product, one must consider the energy levels of transition states in either reaction and their levels of reversibility."
Activation energy	The student refers to the activation energy of the reaction. This can include mentions of the "first activation energy" or "sum of activation energies."	"This indicates that this pathway is more likely to progress because the overall sum of activation energies is lower than for B."
Intermediates	The student describes the energy level for the intermediates of the mechanism.	"Also, in mechanism B, intermediate R is nearly as stable as the desired product and it would take 16 kcal/mol to continue the reaction forward—this means that intermediate R would probably form in high amounts, and it would probably reverse back to the starting material as well."
Kinetics	The student uses the word "kinetics" to describe the kinetics of the reactions.	"Thus, mechanism B is more kinetically favored and more likely to occur."
Rate determining step	The student uses the phrase "rate determining step" to specifically describe a mechanistic step.	"The rate determining step(s) for mechanism A is 15 kcal/mol, compared to 19 kcal/mol in mechanism B."
Rate/time	The student refers to the rate, time, or speed of the reactions. (As the literature indicates students' challenges with interpreting the *x*-axis on RCDs, these features were captured by a single code so as not to interpret unintended meaning in responses referring to the rate, time, or speed of reactions.)	"Mechanism A is preferred because it is selective for the desired keto-aldol product, has lower energy transition states and has a faster overall reaction time."
Energy values	The student uses specific energy values or labels from the reaction coordinate diagram.	"Mechanism A's rate-determining step requires 15 kcal/mol of energy and occurs between the reactant a and intermediate c, while mechanism B's rate-determining step uses 17 kcal/mol of energy and occurs between intermediate r and intermediate t."
General EPF	The student uses general features of the EPF schemes, including descriptions of electron-pushing or changes in bonding, to justify their choice of the most likely mechanism.	"Mechanism A occurs without forming and breaking a bond between the catalyst and reagent and only uses one equivalent of TBD, making this mechanism much more favorable and cost-effective for your company."
Equivalents TBD	The student refers to the equivalents of TBD catalyst added to the reaction.	"Additionally, Pathway A only requires one equivalent of TBD where Pathway B requires two."
TBD adding	The student describes the addition of TBD to the aldehyde in mechanism B and/or specifically uses the words or phrases "attaching," "acting as a nucleophile," "protecting group," or "complexation" to describe the addition of TBD to the aldehyde.	"All the steps seen in mechanism A occur in mechanism B, however, Mechanism B involves additional steps that involve the addition and removal of TBD for the protecting and deprotecting of the aldehyde."
Counting	The student counts transition states, intermediates, or reaction steps OR uses words, such as "fewer," "more," "additional," or "only," referring to the number of steps in each reaction.	"This mechanism has less steps than mechanism A, and therefore the reaction would take less energy to create the product."

By-products	The student refers to the formation of unwanted by-products, usually when students make an argument that more steps lead to more unwanted by-products.	"Mechanism A involves only one intermediate, but with a higher activation energy, while Mechanism B offers a lower activation energy for its first step, but with several more intermediate structures/byproducts."
Reversibility	The student refers to the reversibility or irreversibility of the reaction, describing whether products are likely (or unlikely) to revert back to reactants OR uses words "reversible" or "irreversible."	"In determining the success of either mechanism in producing the desired product, one must consider the energy levels of transition states in either reaction and their levels of reversibility."
Functional group	The student refers to a functional group (commonly ketone, aldehyde, or enol) to make an argument about the reactivity, stability, favourability, or likelihood of one reaction pathway compared to the other.	"The aldehyde is much more reactive than the ketone and therefore would use less energy."
Implied energetics	The student uses the words "stable," "unstable," "favorable," "unfavorable," "more likely," or "less likely" to make an argument about the likelihood of the chosen reaction pathway, without referring specifically to energy levels, features on the energy diagram, or to specific functional groups.	"This intermediate in Mechanism A is more stable than those in B's due to hydrogen bonding and is therefore more favorable."

Table 7.6 Coding scheme for the third analysis stage, in which peer review comments were coded for instances of clear agreement or clear disagreement

Code	Definition	Exemplar
Agree	Within the provided peer review comment, the student indicated clear agreement with the mechanism the author of the reviewed draft indicated as most likely.	"They chose Mechanism A as the pathway that the molecules would most likely go through… I think the author made the correct choice, but there isn't much evidence to back up their claim…"
Disagree	Within the provided peer review comment, the student indicated clear disagreement with the mechanism the author of the reviewed draft indicated as most likely.	"The author chose mechanism B as the preferred method… I personally disagree since more equivalents of TBD were required…"
Neutral	Within the provided peer review comment, the student indicated neither clear agreement nor clear disagreement with the mechanism the author of the reviewed draft indicated as most likely.	"The author chose Mechanism B and it seems reasonable how product formation has become selective with the creation of an alcohol. My explanation was for Mechanism A and how a lower transition state energy would be more easier to overcome and thus at a higher probability rate."

7.11.3 Appendix 3. Examples of coded responses

Sample student responses

Response	Codes
Student 1, Initial draft I believe that the pathway this reaction undergoes is the second. My decision is largely due to the first step of this reaction. The first step of the second pathway is *addition to an aldehyde*, while the first step of the first pathway is an *enolization reaction*. *Aldehydes are very reactive to additions*, while *enolization reactions* are typically more unfavorable. This can be seen in the reaction coordinate diagrams, where the *transition state and product for the first step of the second pathway is lower* than those for the first step of the first pathway. I believe that an *addition to an aldehyde* would be more like to occur than an *enolization reaction* under the same conditions, and therefore the second pathway is observed.	Mechanism B *TBD adding, General EPF* *TBD adding, Functional group* *Transition states* *TBD adding* *General EPF*
Student 1, Revised draft I believe that the pathway this reaction undergoes is pathway A. I believe that this is due to the fact the *rate-determining step* for this reaction is lower in energy than that of pathway B. The first step of pathway B is *more favorable*, so it is likely that that is the first interaction that will take place. However, aldol reactions are *reversible*. The case with *reversible reactions* is that the *most stable mechanism* will occur. Because of the *rate-determining steps*, pathway A is *overall the more stable reaction*, so I believe that pathway A is the most likely to occur. Therefore, I propose that pathway B will not continue beyond the first step, and *it will instead reverse* and allow pathway A to continue and complete the reaction.	Mechanism A *Rate determining step* *Stability/favorability* *Reversibility* *Stability/favorability, Rate determining step* *Stability/favorability* *Reversibility*
Student 2, Initial draft Both of these mechanisms require the same starting materials, follow some of the same steps, and produce the same product and byproduct. The difference between the two is that one reaction pathway will occur more often than the other. It is predicted that, in this situation, Mechanism B would occur more often than Mechanism A. This is due to the fact that, while Mechanism A begins with interactions with a ketone, Mechanism B begins with an interaction with an aldehyde. In general, *aldehydes react faster and easier than ketones*, and would therefore be affected first in this reaction.	Mechanism B *Functional group*
Student 2, Revised draft Both of these mechanisms require the same starting materials, follow some of the same steps, and produce the same product and byproduct. The difference between the two is that one reaction pathway will occur more often than the other. It is predicted that, in this situation, Mechanism A would occur more often than Mechanism B. This is due to the fact that mechanism A's *rate-determining step* requires an *activation energy* of approximately *15 kcal/mol*, while Mechanism B's *rate-determining step* requires a higher *activation energy* of approximately *16 kcal/mol*. Because the *activation energy* (the energy required which will allow the reaction to occur) for Mechanism A is lower, the reaction is easier to "kick-start", and is therefore predicted to occur more often than Mechanism B.	Mechanism A *Rate determining step* *Activation energy, Energy values* *Rate determining step, Activation energy* *Energy values, Activation energy*

Figure 7.9 Two examples of students' initial and revised drafts with the codes applied from the first and second analysis stages.

7.11.4 Appendix 4. Tabular results and statistics

Table 7.7 Mean and standard deviation for each coded feature of students' responses among students selecting either Mechanism A or Mechanism B as most likely. Reported *p*-values are from the Mann–Whitney *U* tests with the Bonferroni correction for multiple hypothesis tests

Code	Selected Mechanism A (*n* = 120), Mean (St. dev.)	Selected Mechanism B (*n* = 44), Mean (St. dev.)	*p*-Values

Thermodynamics	0.08 (0.29)	0.34 (0.86)	0.062
General energy	1.01 (0.89)	0.95 (1.12)	1.000
Transition states	0.62 (0.79)	0.57 (0.97)	1.000
Activation energy	0.90 (1.23)	0.84 (1.27)	1.000
Intermediates	0.55 (0.99)	0.14 (0.51)	0.074
Kinetics	0.12 (0.41)	0.25 (0.72)	1.000
Rate determining step	0.24 (0.74)	0.07 (0.25)	1.000
Rate/time	0.28 (0.62)	0.27 (0.66)	1.000
Energy values	0.98 (1.17)	0.91 (1.20)	1.000
General EPF	0.76 (1.17)	1.34 (1.63)	0.878
Equivalents TBD	0.36 (0.58)	0.18 (0.58)	0.310
TBD adding	0.13 (0.47)	0.50 (0.88)	0.006**
Counting	0.97 (0.89)	0.55 (0.66)	0.107
By-products	0.07 (0.28)	0.14 (0.51)	1.000
Reversibility	0.08 (0.32)	0.39 (0.89)	0.077
Functional group	0.18 (0.67)	0.75 (0.99)	<0.001***
Implied energetics	0.09 (0.32)	0.09 (0.29)	1.000
Total number of codes	7.39 (3.67)	8.27 (4.41)	1.000
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$			

Table 7.8 Mean and standard deviation for the changes in each coded feature of students' revisions, among the groups of students by the nature of their global revisions. Reported *p*-values are from the Kruskal–Wallis one-way analysis of variance by ranks tests adjusted with the Bonferroni correction for multiple hypothesis tests

Code	**Change in number of codes for A→A (*n* = 114), Mean (St. dev)**	**Change in number of codes for A→B (*n* = 6), Mean (St. dev)**	**Change in number of codes for B→A (*n* = 16), Mean (St. dev)**	**Change in number of codes for B→B (*n* = 28), Mean (St. dev)**	***p*-Value**
Thermodynamics	0.05 (0.22)	0.00 (0.00)	0.06 (0.25)	0.04 (0.43)	1.000
General energy	0.21 (0.84)	0.00 (0.63)	0.63 (1.15)	0.36 (0.95)	1.000
Transition states	0.11 (0.56)	0.17 (0.98)	0.13 (0.89)	0.00 (0.47)	1.000
Activation energy	0.21 (0.95)	0.17 (0.41)	0.19 (1.17)	0.14 (0.65)	1.000
Intermediates	-0.02 (0.44)	-0.33 (1.03)	0.25 (0.68)	0.18 (0.55)	1.000
Kinetics	0.07 (0.34)	-0.17 (0.41)	0.00 (0.00)	0.07 (0.38)	1.000
Rate determining step	0.18 (0.57)	0.17 (0.41)	0.38 (0.62)	0.00 (0.00)	0.697
Rate/time	0.05 (0.35)	-0.67 (1.21)	0.06 (0.68)	0.04 (0.19)	1.000
Energy values	0.27 (0.67)	0.17 (0.41)	0.25 (1.13)	0.32 (0.94)	1.000
General EPF	0.04 (0.40)	0.33 (0.52)	-0.50 (1.26)	0.43 (1.26)	1.000
Equivalents TBD	0.16 (0.43)	0.00 (0.63)	0.13 (0.81)	-0.07 (0.26)	1.000
TBD adding	0.09 (0.37)	1.17 (1.17)	-0.06 (1.00)	0.25 (0.75)	0.003 **
Counting	0.04 (0.73)	-0.33 (0.52)	0.63 (1.02)	0.11 (0.42)	0.028 *
By-products	0.01 (0.16)	0.00 (0.00)	0.00 (0.00)	0.18 (0.94)	1.000
Reversibility	0.04 (0.24)	-0.17 (0.41)	0.31 (0.87)	0.07 (0.72)	1.000
Functional group	0.02 (0.13)	0.83 (1.33)	-0.44 (0.81)	0.32 (0.82)	< 0.001 ***
Implied energetics	0.01 (0.31)	0.00 (0.00)	0.25 (0.77)	0.00 (0.27)	1.000
Total number of codes	1.54 (2.50)	1.33 (3.14)	2.25 (3.42)	2.43 (3.17)	1.000
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$					

Table 7.9 The p-values for the *post hoc* pairwise Mann–Whitney U tests for statistically significant codes in the Kruskal–Wallis tests. The p-values are adjusted using the Bonferroni correction for multiple hypothesis tests

Code		A→A	A→B	B→A
TBD adding	A→B	< 0.001 ***	—	—
	B→A	1.000	0.099	—
	B→B	1.000	0.101	1.000
Counting	A→B	0.520	—	—
	B→A	0.008 **	0.124	—
	B→B	1.000	0.066	0.018 *
Functional group	A→B	< 0.001 ***	—	—
	B→A	< 0.001 ***	0.205	—
	B→B	0.009 **	1.000	0.031 *
$*p < 0.05$, $**p < 0.01$, $***p < 0.001$				

Table 7.10 Descriptive statistics for variables included in the logistic regression analysis

Variable	Mean	St. dev.	Min	Max
Initial draft mech. ($A = 0$, $B = 1$)	0.25	0.43	0	1
Comments received (freq. of disagreement)	0.57	0.79	0	4
Drafts reviewed (freq. of disagreement)	1.15	1.04	0	3

7.12 Acknowledgements

The authors would like to thank the Keck Foundation and the University of Michigan Third Century Initiative for funding. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260. The authors would additionally like to thank Solaire A. Finkenstaedt-Quinn, Jennifer A. Schmidt-McCormack, Ina Zaimi, Ashley Karlin, Atia Sattar, Barry C. Thompson, and Anne Ruggles Gere for their input on the WTL assignment central to this investigation. Furthermore, we would like to thank Amber Dood, Eleni Zotos, and members of the Shultz group for discussions related to the preparation of this manuscript.

7.13 References

(1) Graulich, N. The Tip of the Iceberg in Organic Chemistry Classes: How Do Students Deal with the Invisible? *Chem. Educ. Res. Pract.* **2015**, *16* (1), 9–21. https://doi.org/10.1039/c4rp00165f.

(2) Lamichhane, R.; Reck, C.; Maltese, A. V. Undergraduate Chemistry Students' Misconceptions about Reaction Coordinate Diagrams. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 834–845. https://doi.org/10.1039/c8rp00045j.

(3) Popova, M.; Bretz, S. L. "It's Only the Major Product That We Care about in Organic Chemistry": An Analysis of Students' Annotations of Reaction Coordinate Diagrams. *J. Chem. Educ.* **2018**, *95* (7), 1086–1093. https://doi.org/10.1021/acs.jchemed.8b00153.

(4) Popova, M.; Bretz, S. L. Organic Chemistry Students' Interpretations of the Surface Features of Reaction Coordinate Diagrams. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 919–931. https://doi.org/10.1039/c8rp00063h.

(5) Popova, M.; Bretz, S. L. Organic Chemistry Students' Challenges with Coherence Formation between Reactions and Reaction Coordinate Diagrams. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 732–745. https://doi.org/10.1039/c8rp00064f.

(6) NGSS Lead States. *Next Generation Science Standards: For States, By States*; The National Academies Press: Washington, 2013.

(7) Passmore, C.; Gouvea, J. S.; Giere, R. Models in Science and in Learning Science: Focusing Scientifi c Practice on Sense-Making. In *International Handbook of Research in History, Philosophy and Science Teaching*; Matthews, M. R., Ed.; Springer Scinece+Business Media: Dordrecht, 2014; pp 1171–1202. https://doi.org/10.1007/978-94-007-7654-8.

(8) Passmore, C.; Schwarz, C. V; Mankowski, J. Developing and Using Models. In *Helping Students Make Sense of the World Using Next Generation Science and Engineering Practices*; Schwarz, C., Passmore, C., Reiser, B. J., Eds.; NSTA Press, 2016; pp 109–134.

(9) Gouvea, J.; Passmore, C. 'Models of' versus 'Models for': Toward an Agent-Based Conception of Modeling in the Science Classroom. *Sci. Educ.* **2017**, *26* (1–2), 49–63. https://doi.org/10.1007/s11191-017-9884-4.

(10) Johnstone, A. H. Why Is Science Difficult to Learn? Things Are Seldom What They Seem. *J. Comput. Assist. Learn.* **1991**, *7*, 75–83.

(11) Taber, K. S. Revisiting the Chemistry Triplet: Drawing upon the Nature of Chemical Knowledge and the Psychology of Learning to Inform Chemistry Education. *Chem. Educ. Res. Pract.* **2013**, *14* (2), 156–168. https://doi.org/10.1039/c3rp00012e.

(12) Goodwin, W. Mechanisms and Chemical Reaction. In *Philosophy of Chemistry*; Hendry, R. F., Needham, P., Woody, A. I., Eds.; Elsevier BV, 2012; Vol. 6, pp 309–327. https://doi.org/10.1016/B978-0-444-51675-6.50023-2.

(13) Kozma, R.; Chin, E.; Russell, J.; Marx, N. The Roles of Representations and Tools in the Chemistry Laboratory and Their Implications for Chemistry Learning. *J. Learn. Sci.* **2000**, *9* (2), 105–143. https://doi.org/10.1207/s15327809jls0902_1.

(14) Kozma, R.; Russell, J. Students Becoming Chemists: Developing Representationl Competence. *Vis. Sci. Educ.* **2005**, 121–145. https://doi.org/10.1007/1-4020-3613-2_8.

(15) Popova, M.; Jones, T. Chemistry Instructors' Intentions toward Developing, Teaching, and Assessing Student Representational Competence Skills. *Chem. Educ. Res. Pract.* **2021**. https://doi.org/10.1039/d0rp00329h.

(16) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Ideas about Nucleophiles and Electrophiles: The Role of Charges and Mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797–810. https://doi.org/10.1039/c5rp00113g.

(17) Galloway, K. R.; Stoyanovich, C.; Flynn, A. B. Students' Interpretations of Mechanistic Language in Organic Chemistry before Learning Reactions. *Chem. Educ. Res. Pract.* **2017**, *18* (2), 353–374. https://doi.org/10.1039/c6rp00231e.

(18) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/c8rp00131f.

(19) Watts, F. M.; Schmidt-McCormack, J.; Wilhelm, C.; Karlin, A.; Sattar, A.; Thompson, B.; Gere, A. R.; Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21*, 1148–1172. https://doi.org/10.1039/C9RP00185A.

(20) Ferguson, R.; Bodner, G. M. Making Sense of the Arrow-Pushing Formalism among Chemistry Majors Enrolled in Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 102–113. https://doi.org/10.1039/b806225k.

(21) Finkenstaedt-Quinn, S. A.; Watts, F. M.; Petterson, M. N.; Archer, S. R.; Snyder-White, E. P.; Shultz, G. V. Exploring Student Thinking about Addition Reactions. *J. Chem. Educ.* **2020**, *97* (7), 1852–1862. https://doi.org/10.1021/acs.jchemed.0c00141.

(22) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.; Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions via App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 878–892. https://doi.org/10.1039/c9rp00260j.

(23) 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.

(24) Cruz-Ramírez De Arellano, D.; Towns, M. H. Students' Understanding of Alkyl Halide Reactions in Undergraduate Organic Chemistry. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 501–515. https://doi.org/10.1039/c3rp00089c.

(25) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Fragmented Ideas about the Structure and Function of Nucleophiles and Electrophiles: A Concept Map Analysis. *Chem.*

Educ. Res. Pract. **2016**, *17* (4), 1019–1029. https://doi.org/10.1039/c6rp00111d.

(26) Dood, A. J.; Dood, J. C.; Cruz-Ramírez De Arellano, D.; Fields, K. B.; Raker, J. R. Analyzing Explanations of Substitution Reactions Using Lexical Analysis and Logistic Regression Techniques. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 267–286. https://doi.org/10.1039/c9rp00148d.

(27) 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.

(28) Graulich, N.; Bhattacharyya, G. Investigating Students' Similarity Judgments in Organic Chemistry. *Chem. Educ. Res. Pract.* **2017**, *18* (4), 774–784. https://doi.org/10.1039/c7rp00055c.

(29) Atkinson, M. B.; Popova, M.; Croisant, M.; Reed, D. J.; Bretz, S. L. Development of the Reaction Coordinate Diagram Inventory: Measuring Student Thinking and Confidence. *J. Chem. Educ.* **2020**, *97* (7), 1841–1851. https://doi.org/10.1021/acs.jchemed.9b01186.

(30) Atkinson, M. B.; Bretz, S. L. Measuring Changes in Undergraduate Chemistry Students' Reasoning with Reaction Coordinate Diagrams: A Longitudinal, Multi-Institution Study. *J. Chem. Educ.* **2021**, *98* (4), 1064–1076. https://doi.org/10.1021/acs.jchemed.0c01419.

(31) Atkinson, M. B.; Croisant, M.; Bretz, S. L. Investigating First-Year Undergraduate Chemistry Students' Reasoning with Reaction Coordinate Diagrams When Choosing among Particulate-Level Reaction Mechanisms. *Chem. Educ. Res. Pract.* **2021**, *22* (1), 199–213. https://doi.org/10.1039/d0rp00193g.

(32) Parobek, A. P.; Chaffin, P. M.; Towns, M. H. Location-Thinking, Value-Thinking, and Graphical Forms: Combining Analytical Frameworks to Analyze Inferences Made by Students When Interpreting the Points and Trends on a Reaction Coordinate Diagram. *Chem. Educ. Res. Pract.* **2021**. https://doi.org/10.1039/d1rp00037c.

(33) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable Reasoning: Multi-Variate Problem-Solving in Organic Chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281–292. https://doi.org/10.1039/c0rp90003f.

(34) Christian, K.; Talanquer, V. Modes of Reasoning in Self-Initiated Study Groups in Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 286–295. https://doi.org/10.1039/c2rp20010d.

(35) 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.

(36) Weinrich, M. L.; Talanquer, V. Mapping Students' Modes of Reasoning When Thinking

about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 394–406. https://doi.org/10.1039/c5rp00208g.

(37) Lieber, L.; Graulich, N. Investigating Students' Argumentation When Judging the Plausibility of Alternative Reaction Pathways in Organic Chemistry. *Chem. Educ. Res. Pract.* **2022**. https://doi.org/10.1039/d1rp00145k.

(38) Grove, N. P.; Bretz, S. L. A Continuum of Learning: From Rote Memorization to Meaningful Learning in Organic Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 201–208. https://doi.org/10.1039/c1rp90069b.

(39) Caspari, I.; Graulich, N. Scaffolding the Structure of Organic Chemistry Students' Multivariate Comparative Mechanistic Reasoning. *Int. J. Phys. Chem. Educ.* **2019**, *11* (2), 31–43. https://doi.org/10.12973/ijpce/211359.

(40) Watts, F. M.; Zaimi, I.; Kranz, D.; Graulich, N.; Shultz, G. V. Investigating Students' Reasoning over Time for Case Comparisons of Acyl Transfer Reaction Mechanisms. *Chem. Educ. Res. Pract.* **2021**, *22*, 364–381. https://doi.org/10.1039/d0rp00298d.

(41) Brandfonbrener, P. B.; Watts, F. M.; Shultz, G. V. Organic Chemistry Students' Written Descriptions and Explanations of Resonance and Its Influence on Reactivity. *J. Chem. Educ.* **2021**. https://doi.org/10.1021/acs.jchemed.1c00660.

(42) Grimberg, B. I.; Hand, B. Cognitive Pathways: Analysis of Students' Written Texts for Science Understanding. *Int. J. Sci. Educ.* **2009**, *31* (4), 503–521. https://doi.org/10.1080/09500690701704805.

(43) Moon, A.; Moeller, R.; Gere, A. R.; Shultz, G. V. Application and Testing of a Framework for Characterizing the Quality of Scientific Reasoning in Chemistry Students' Writing on Ocean Acidification. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 484–494. https://doi.org/10.1039/c9rp00005d.

(44) Moreira, P.; Marzabal, A.; Talanquer, V. Using a Mechanistic Framework to Characterise Chemistry Students' Reasoning in Written Explanations. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 120–131. https://doi.org/10.1039/c8rp00159f.

(45) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Res. Teach. English* **2015**, *50* (2), 199–235.

(46) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and Conceptual Learning in Science: An Analysis of Assignments. *Writ. Commun.* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(47) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Chambers, T. G.; Moon, A.; Goldman, R. S.; Gere, A. R.; Shultz, G. V. Investigation of the Influence of a Writing-To-Learn Assignment on

Student Understanding of Polymer Properties. *J. Chem. Educ.* **2017**, *94* (11), 1610–1617. https://doi.org/10.1021/acs.jchemed.7b00363.

(48) Halim, A. S.; Finkenstaedt-Quinn, S. A.; Olsen, L. J.; Gere, A. R.; Shultz, G. V. Identifying and Remediating Student Misconceptions in Introductory Biology via Writing-to-Learn Assignments and Peer Review. *CBE Life Sci. Educ.* **2018**, *17* (2), 1–12. https://doi.org/10.1187/cbe.17-10-0212.

(49) Moon, A.; Zotos, E.; Finkenstaedt-Quinn, S.; Gere, A. R.; Shultz, G. Investigation of the Role of Writing-to-Learn in Promoting Student Understanding of Light-Matter Interactions. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 807–818. https://doi.org/10.1039/c8rp00090e.

(50) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(51) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Kasner, G.; Wilhelm, C. A.; Moon, A.; Gere, A. R.; Shultz, G. V. Capturing Student Conceptions of Thermodynamics and Kinetics Using Writing. *Chem. Educ. Res. Pract.* **2020**, *21*, 922–939. https://doi.org/10.1039/C9RP00292H.

(52) Marks, L.; Lu, H.; Chambers, T.; Finkenstaedt-Quinn, S.; Goldman, R. S. Writing-to-Learn in Introductory Materials Science and Engineering. *MRS Commun.* **2022**, *XX* (xx), 1–11. https://doi.org/10.1557/s43579-021-00114-z.

(53) Finkenstaedt-Quinn, S. A.; Polakowski, N.; Gunderson, B.; Shultz, G. V.; Gere, A. R. Utilizing Peer Review and Revision in STEM to Support the Development of Conceptual Knowledge Through Writing. *Writ. Commun.* **2021**. https://doi.org/10.1177/07410883211006038.

(54) Finkenstaedt-Quinn, S. A.; Petterson, M.; Gere, A.; Shultz, G. Praxis of Writing-to-Learn: A Model for the Design and Propagation of Writing-to-Learn in STEM. *J. Chem. Educ.* **2021**, *98* (5), 1548–1555. https://doi.org/10.1021/acs.jchemed.0c01482.

(55) Finkenstaedt-Quinn, S. A.; Snyder-White, E. P.; Connor, M. C.; Gere, A. R.; Shultz, G. V. Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *J. Chem. Educ.* **2019**, *96* (2), 227–237. https://doi.org/10.1021/acs.jchemed.8b00711.

(56) Lundstrom, K.; Baker, W. To Give Is Better than to Receive: The Benefits of Peer Review to the Reviewer's Own Writing. *J. Second Lang. Writ.* **2009**, *18* (1), 30–43. https://doi.org/10.1016/j.jslw.2008.06.002.

(57) Cho, K.; MacArthur, C. Student Revision with Peer and Expert Reviewing. *Learn. Instr.* **2010**, *20* (4), 328–338. https://doi.org/10.1016/j.learninstruc.2009.08.006.

(58) Cho, Y. H.; Cho, K. Peer Reviewers Learn from Giving Comments. *Instr. Sci.* **2011**, *39* (5), 629–643. https://doi.org/10.1007/s11251-010-9146-1.

(59) Berg, S. A.; Moon, A. Prompting Hypothetical Social Comparisons to Support Chemistry Students' Data Analysis and Interpretations. *Chem. Educ. Res. Pract.* **2022**. https://doi.org/10.1039/d1rp00213a.

(60) Gupte, T.; Watts, F. M.; Schmidt-McCormack, J. A.; Zaimi, I.; Gere, A. R.; Shultz, G. V. Students' Meaningful Learning Experiences from Participating in Organic Chemistry Writing-to-Learn Activities. *Chem. Educ. Res. Pract.* **2021**, *22*, 396–414. https://doi.org/10.1039/d0rp00266f.

(61) Petterson, M. N.; Finkenstaedt-Quinn, S. A.; Gere, A. R.; Shultz, G. V. The Role of Authentic Contexts and Social Elements in Supporting Organic Chemistry Students' Interactions with Writing-to-Learn Assignments. *Chem. Educ. Res. Pract.* **2022**. https://doi.org/10.1039/d1rp00181g.

(62) Kozma, R. B.; Russell, J. Multimedia and Understanding: Expert and Novice Responses to Different Representations of Chemical Phenomena. *J. Res. Sci. Teach.* **1997**, *34* (9), 949–968. https://doi.org/10.1002/(sici)1098-2736(199711)34:9<949::aid-tea7>3.3.co;2-f.

(63) Raker, J. R.; Holme, T. A.; Murphy, K. L. The ACS Exams Institute Undergraduate Chemistry Anchoring Concepts Content Map IV: Physical Chemistry. *J. Chem. Educ.* **2013**, *95* (2), 238–241. https://doi.org/10.1021/acs.jchemed.7b00531.

(64) Flower, L.; Hayes, J. R. A Cognitive Process Theory of Writing. *Coll. Compos. Commun.* **1981**, *32* (4), 365–387.

(65) Flower, L.; Hayes, J. R. Images, Plans, and Prose: The Representation of Meaning in Writing. *Writ. Commun.* **1984**, *1* (1), 120–160.

(66) Hayes, J. R. A New Framework for Understanding Cognition and Affect in Writing. In *The Science of Writing: Theories, Methods, Individual Differences, and Applications*; Levy, C. M., Ransdell, S., Eds.; Lawrence Erbaum Associates: Mahwah, New Jersey, 1996; pp 1–27.

(67) Hammer, D.; Elby, A. Tapping Epistemological Resources for Learning Physics. *J. Learn. Sci.* **2003**, *12* (1), 53–90. https://doi.org/10.1207/S15327809JLS1201.

(68) Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F. Resources, Framing, and Transfer. In *Transfer of Learning: Research and Perspectives*; Mestre, J., Ed.; Information Age Publishing: Greenwich, CT, 2004.

(69) Hammar, P.; Ghobril, C.; Antheaume, C.; Wagner, A.; Baati, R.; Himo, F. Theoretical Mechanistic Study of the TBD-Catalyzed Intramolecular Aldol Reaction of Ketoaldehydes.

J. Org. Chem. **2010**, *75* (14), 4728–4736. https://doi.org/10.1021/jo100488g.

(70) Miles, M. B.; Huberman, A. M.; Saldana, J. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed.; Sage: Los Angeles, CA, 2014.

(71) Greene, J. C. Is Mixed Methods Social Inquiry a Distinctive Methodology? *J. Mix. Methods Res.* **2008**, *2* (1), 7–22. https://doi.org/10.1177/1558689807309969.

(72) QSR International Pty Ltd. NVivo Qualitative Data Analysis Software (Version 12). 2018.

(73) Watts, F. M.; Finkenstaedt-Quinn, S. A. The Current State of Methods for Establishing Reliability in Qualitative Chemistry Education Research Articles. *Chem. Educ. Res. Pract.* **2021**, No. 22, 565–578. https://doi.org/10.1039/d1rp00007a.

(74) Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20* (1), 37–46.

(75) Kirilenko, A. P.; Stepchenkova, S. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLoS One* **2016**, *11* (3), 1–14. https://doi.org/10.1371/journal.pone.0149787.

(76) RStudio Team. RStudio: Integrated Development for R. 2018.

(77) Sheskin, D. J. *Handbook of Parametric and Nonparametric Statistical Procedures*, 5th ed.; Press, CRC: Boca Raton, 2011.

(78) Jaccard, J. Interactions Between Qualitative and Quantitative / Continuous Predictors. In *Interaction Effects in Logistic Regression*; SAGE Publications, Inc.: Thousand Oaks, 2011; pp 31–41.

(79) Anderson, T. L.; Bodner, G. M. What Can We Do about "Parker"? A Case Study of a Good Student Who Didn't "get" Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–101. https://doi.org/10.1039/b806223b.

Chapter 8

Investigating Writing-To-Learn To Elicit Organic Chemistry Students' Mechanistic Reasoning

8.1 Initial remarks

This chapter is the final of three chapters which involve analyzing student responses to different writing-to-learn (WTL) assignments to examine how WTL can elicit students' reasoning in organic chemistry. This chapter follows from Chapter 7 by presenting another study that examines students' reasoning with reaction mechanisms in organic chemistry. Specifically, this chapter examines a WTL assignment that elicits students' mechanistic reasoning for a single organic chemistry reaction by asking students to explain the mechanism for the formation of two products in an acid-catalyzed amide hydrolysis reaction. This chapter presents an analytical framework derived from the philosophy of science literature for recognizing features of students' mechanistic reasoning. The analysis in this chapter serves as the basis for the remaining chapters of the dissertation, which extend the analytical framework to develop machine learning models for automatically analyzing student writing to deliver tailored, formative feedback.

Using the cognitive process theory of writing to ground the analysis of student responses to the WTL assignment, this study contributes further evidence that examining the artifacts of student written responses is one approach for making inferences about students' reasoning. The results from this study provide a qualitative description of the variety of ways in which students included features necessary for mechanistic reasoning in their writing. Specifically, the results present how students described an overview of the phenomenon, the conditions necessary for the mechanism to occur, the changes that take place during the mechanism, and the properties of reacting species which drive the changes that occur. For example, students described the changes that take place during the mechanism in a variety of ways: describing electron movement explicitly (e.g., using the words "electron" or "lone pairs"), describing electron movement implicitly (e.g., using the words "attacks" or "protonates"), and describing the actual changes in bonding that occur. Understanding the different ways students describe these changes is valuable, because the research literature indicates that students do not always consider the underlying electronic movement inherent to reaction mechanisms. Additionally, the study involves the use of an association metric, called lift, to measure the degree of co-occurrence between different features within students' responses. Using this measure provided evidence for the hierarchical nature of students' mechanistic descriptions, in that the features of mechanistic reasoning within students' writing tended to be more associated with features of similar levels of complexity. For example, students' descriptions of the overview of the phenomenon (the lowest level of the framework) overlapped most with student's descriptions of the conditions necessary for the mechanism to occur (the next level of the framework). Additionally, the lift metric provided evidence that students were making appropriate connections between mechanistic steps and the properties of reacting species, such as using acid–base reasoning when discussing protonation or deprotonation steps. Overall, the findings indicate the capacity for analyzing writing to make inferences about students' mechanistic reasoning while providing implications for instructors seeking to support students in making connections between reaction mechanisms and the underlying chemical reasoning.

This chapter was originally published as a research article in *Chemistry Education Research and Practice*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author, I contributed to conceptualization, methodology, data analysis, and writing (both original draft preparation and review and editing). J.A. Schmidt-McCormack contributed to conceptualization, data collection, and data analysis, and C.A. Wilhelm contributed to data analysis. A. Karlin and A. Sattar contributed to conceptualization and data collection. B.C. Thompson contributed to funding acquisition, conceptualization, and data analysis. A.R. Gere, and G.V. Shultz contributed to funding acquisition, conceptualization, data collection, and writing (review and editing).

<u>*Original publication and copyright information:*</u>

8.2 Abstract

Learning to reason through organic reaction mechanisms is challenging for students because of the volume of reactions covered in introductory organic chemistry and the complexity of conceptual knowledge and reasoning skills required to develop meaningful understanding. However, understanding reaction mechanisms is valuable for students because they are useful for predicting and explaining reaction outcomes. To identify the features students find pertinent when explaining reaction mechanisms, we have collected students' written descriptions of an acid-catalyzed amide hydrolysis reaction. Students' writing was produced during the implementation of writing-to-learn assignments in a second semester organic chemistry laboratory course. We analyzed students' written responses using an analytical framework for recognizing students' mechanistic reasoning, originally developed with attention to the philosophy of science literature. The analysis sought to identify the presence of specific features necessary for mechanistic reasoning belonging to four broad categories: (1) describing an overview of the reaction, (2) detailing the setup conditions required for the mechanism to occur, (3) describing the changes that take place over the course of the mechanism, and (4) identifying the properties of reacting species. This work provides a qualitative description of the variety of ways in which students included these features necessary for mechanistic reasoning in their writing. We additionally analyzed instances of co-occurrence for these features in students' writing to make inferences about students' mechanistic reasoning, defined here as the use of chemical properties to justify how electrons, atoms, and molecules are reorganized over the course of a reaction. Feature co-occurrences were quantified using the lift metric to measure the degree of their mutual dependence. The quantitative lift results provide empirical support for the hierarchical nature of students' mechanistic descriptions and indicate the variation in students' descriptions of mechanistic change in conjunction with appeals to chemistry concepts. This research applies a framework for identifying the features present in students' written mechanistic descriptions, and illustrates the use of an association metric to make inferences about students' mechanistic reasoning. The findings reveal the capacity of implementing and analyzing writing to make inferences about students' mechanistic reasoning.

8.3 Introduction

Organic chemistry is a challenging subject, largely because of the volume of reaction mechanisms presented in the course, which are especially difficult for students to learn meaningfully. This challenge is due in part to the conceptual nature of the discipline[1,2] and is related to the types of problem solving skills required for success in the organic chemistry classroom.[3,4] Previous research has focused on this acknowledged difficulty, including investigations characterizing the use and usefulness of the electron-pushing formalism;[5–8] research examining students' use of conceptual reasoning applied to reaction mechanisms;[9–13] and studies involving restructuring the curricula for general chemistry[14] or organic chemistry[15–19] to promote students' understanding of the connections between chemical structure, properties, and reactivity.

Understanding how students both describe and explain reaction mechanisms is valuable because of the inherent challenge of learning to use the electron-pushing formalism while connecting steps in a mechanism to conceptual understanding. A means to access students' descriptions and explanations on a large scale is through students' writing. Writing-to-learn (WTL) is a pedagogical practice that instructs students to produce written artefacts of their knowledge, which can serve as a resource for understanding students' reasoning[20–22] while serving to promote students' conceptual understanding.[23–28]

The goal of this study is to investigate the mechanistic reasoning used by a large number of students by analyzing their written responses to a WTL prompt meant to elicit mechanistic reasoning about a specific reaction mechanism. The first objective of the analysis is to describe the variations in the way students write about the components they found pertinent when describing and explaining the mechanism, coded as features necessary for engaging in mechanistic reasoning. The second objective of the analysis is to identify students' engagement in mechanistic reasoning by examining the co-occurrences of these features. Note that, although there is no consensus on the definition of mechanistic reasoning,[29] for the purposes of this study, we conceptualize mechanistic reasoning as the ability to identify the species involved over the course of a reaction (e.g., the starting materials, intermediates, and products), to provide an account for how molecules change over the course of a reaction, and to appeal to chemical properties to justify why these changes occur. This definition aligns with the common features present in the various definitions of mechanistic reasoning identified by organic chemistry faculty,[29] and this definition aligns with those identified in prior studies.[10,21,30,31] In particular, this definition of mechanistic reasoning requires both the *what* and *how* for a reaction—i.e., describing *what* structural changes occur from starting materials to intermediates to products and *how* these changes arise from interactions between the involved subcomponents (electrons, atoms, and molecules). This definition also requires justifications for *why* mechanistic steps occur by appealing to the properties of involved components (e.g., nucleophilicity and electrophilicity). Note that this definition of mechanistic reasoning is distinct from some definitions of causal mechanistic reasoning, which also require an energetic justification for why a reaction proceeds as it does from one step to the next.[13,32]

8.3.1 Mechanistic reasoning in organic chemistry

Mechanisms are used by organic chemists to explain or predict the outcome of reactions. Because of their usefulness, the organic chemistry curriculum typically involves a study of the mechanisms for each class of reaction presented to students, and problems are often posed assuming students will be able to use mechanisms as a problem-solving tool[7,8] Hence, the ability to reason through a reaction mechanism is a useful skill that can help students achieve success in organic chemistry.[7]

However, research has shown that many students do not use mechanisms meaningfully and that students often do not value the electron-pushing formalism in the same way as practicing chemists.[7,8] Additionally, studies found that students may not conceptualize the electron-pushing formalism to have any physical meaning,[5,6] though this was shown not to be true in a modified curriculum.[18,19] Prior research also suggests that students hold a range of intuitions, misconceptions, and understandings regarding fundamental concepts pertaining to organic reaction mechanisms.[10,12,33–35] Although students might have some conceptual understanding—and are often able to produce correct mechanisms for common reactions—studies have demonstrated that they often lack the ability to connect chemical reasoning to individual steps in a reaction mechanism.[3–6]

Particular barriers to students' learning are their approaches to problem-solving, which may be either product- or process-oriented. Product-oriented approaches incorporate reasoning focused on the final product, result, or answer to the problem rather than the process or methods by which the solution is obtained. Process-oriented approaches include model-based reasoning, in which mechanistic explanations are developed using generalized mental models about structure and reactivity,[3,36] and are reflected in students' use of causal or multi-component argumentation to explain chemical reactions.[10,31,37,38] Successful process-oriented approaches also include reasoning that demonstrates knowledge of the connections between properties of reacting species (e.g., basicity or nucleophilicity) and the mechanistic steps of a reaction.[39] Process-oriented problem-solving requires students to reason about the process of a reaction as opposed to reasoning only about the reactants and products. This type of problem-solving values the usefulness of mechanisms to explain or predict reaction outcomes, and is hence an important skill to develop when learning organic chemistry.[4]

Despite the importance of the process of a mechanism, students often engage in product-oriented problem-solving.[4] This type of problem-solving is evident in students' drawn mechanisms which often demonstrate a focus on simply illustrating mechanistic steps to arrive at the given product without considering whether or not the steps shown are chemically reasonable.[5,12,32] Product-oriented strategies include reasoning based on remembered cases or rules that are prompted by the surface features of molecules,[3,36,39] and are evident in studies demonstrating students' use of descriptive or relational argumentation that lacks consideration of multiple components or cause–effect relationships when explaining chemical reactions.[10,31,37,38] Additionally, product-oriented strategies are evident in studies illustrating that students do not necessarily consider alternative reaction pathways or the dynamic, rather than static, nature of chemical reactions.[13,40] A possible reason that students focus on product- rather than process-oriented problem solving is that general chemistry tends to foster product-oriented strategies, so many of the problem-solving skills students have learned in prior courses do not transfer to organic chemistry.[1,2]

The disciplinary skills and conceptual knowledge with which students must be proficient while solving mechanistic problems is an additional barrier to learning. Students must have representational competence, and they must engage with many concepts fundamental to understanding mechanisms, including recognizing reactants as acids and bases or as nucleophiles and electrophiles.[4] Because students must access many types of information when working with mechanisms, it can be difficult for them to make connections between what occurs in a mechanism and the chemical explanations underlying each step. This issue of cognitive load has been suggested to contribute to students' devaluation of mechanisms for problem-solving purposes[7] and is connected to the concern that mechanisms are usually taught in a way that encourages memorization (a product-oriented approach) and discourages chemical understanding (a process-oriented approach).[18] The research in mechanistic reasoning has identified students' struggles with learning mechanisms, detailing how students solve problems or explain reactions with a focus on the answer rather than using chemical reasoning to understand the process. The literature demonstrates that this lack of engagement is connected to problems of cognitive load and lack of sophisticated chemical understanding. These findings provide space for research-based instructional practices that promote students' abilities to apply chemical reasoning to reaction mechanisms.

8.3.2 Using writing-to-learn to access students' mechanistic reasoning

An instructional practice that requires students to engage with mechanisms beyond working with the electron-pushing formalism is writing-to-learn (WTL), which involves using writing assignments to engage students with course content. The primary goal of WTL is to foster students' deeper conceptual understanding.[27,41] WTL has been implemented in the context of chemistry courses and has been shown to support development of conceptual knowledge and disciplinary reasoning skills.[20,22,24–26,28,42,43]

WTL can be leveraged in the context of organic chemistry to help students identify the value in utilizing mechanisms to solve problems. Using WTL in this way is motivated by the idea that writing offers a valuable route into the electron-pushing formalism, which prior researchers recognized as a language that students must first learn and understand before being able to use successfully when engaging in reasoning.[7,16–18] As opposed to problems requiring students to use the electron-pushing formalism—problems which assume that students will implicitly make connections between mechanistic representations and chemical reasoning—writing requires students to explicitly make such connections. This allows researchers to use students' writing to infer and analyze their reasoning, and for the work of many students to be analyzed (as opposed to interview analysis which is typically limited to a small subset of students).

8.4 Theoretical framework

This research is grounded in theories of writing as a tool for learning, with particular attention to perspectives on the cognitive processes that occur during writing.[44–46] These theories not only justify the implementation of WTL pedagogies,[45,47] but also serve as a theoretical basis for analyzing students' written work for evidence of mechanistic reasoning. This study is specifically guided by the cognitive process theory of writing originally proposed by Flower and Hayes[48,49] and later revised by Hayes.[50] This theory states that learning occurs when writers must access content knowledge and address content problems to meet their writing goals. Components of the theory include the social environment, the motivation for writing, and the cognitive moves that are made while writing.[50] The theory identifies three cognitive processes—planning, writing, and revising—that occur at every point during the production of a text. These processes occur in the context of the task environment—including the problem or prompt, the text-in-production, and the social environment—and require the writer to access any available knowledge of the topic.[48] During these processes, the writer must form internal representations of knowledge, translate these representations into language, and evaluate and revise the text being written.[49] This is where learning can occur, as the writer must explore and consolidate knowledge for the purpose of translating representations into written language.

The cognitive process theory of writing provides ground for utilizing students' written work as an analytical tool for understanding students' knowledge. Writing a mechanistic description requires students to find or produce the symbolically represented reaction mechanism and to translate it into words, using their knowledge of fundamental chemistry concepts to explain why mechanistic steps occur. While doing this translation, students engage in the recursive process of writing which requires them to explore their knowledge and revisit their ideas. While there is a possibility that students might use appropriate jargon without actually understanding the language they are using,[6] the cognitive process theory posits that when using these words in their writing, students are at least engaging with the related concepts. The analysis of students' writing relies on the fact that students are given time to decide what information to include and not include. Thus, when a student chooses to include (or, during the process of writing, does not include) some aspect necessary to engage in reasoning, it can provide insight into what content students do and do not find relevant when explaining a reaction mechanism. For these reasons, students' writing can serve as a useful source of data for understanding students' reasoning.

8.5 Research questions

The present study examines students' responses to a writing assignment eliciting descriptions of an organic reaction mechanism. The research seeks to address the following questions to demonstrate the use of writing analysis to make inferences about students' mechanistic reasoning:

1. What features necessary for mechanistic reasoning are present in students' written descriptions of an organic reaction mechanism?
2. How do students write about each feature?
3. What inferences about students' mechanistic reasoning can be made by analyzing co-occurrences of the features necessary for mechanistic reasoning?

8.6 Methods

8.6.1 Setting and participants

The study was conducted at a large, Midwestern research university within a second-semester organic chemistry laboratory course (often taken concurrently with the second-semester lecture course). The laboratory course includes a lecture and laboratory component, both of which meet once a week. The lecture is taught by faculty and postdoctoral instructors who describe experiments and procedures, and the laboratory is facilitated by graduate teaching assistants. The coursework requires students to maintain a laboratory notebook, complete three writing assignments (one of which is the focus of this study), and take quizzes for assessment. The three writing assignments made up thirty percent of students' grades, with each writing assignment contributing ten percent. The participants consisted of the 543 students who received a final score in the course and completed the WTL assignment described below.

8.6.2 Writing-to-learn assignment

The WTL assignment was the third and final WTL assignment that students completed during the semester. It was developed in collaboration with researchers experienced in designing writing assignments to support meaningful learning and with attention to components of the cognitive process theory of writing.[27,50] The relevant prompt components are specified in Figure 8.1, with the full prompt reproduced in 8.11.1 Appendix 1. The prompt design included consideration of components meant to elicit mechanistic reasoning by describing that thalidomide undergoes acid-catalyzed hydrolysis and explicitly illustrating two hydrolysis products. Students were asked to describe the mechanism for the formation of both hydrolysis products and to propose an analog that would prevent the mechanism. For reference, one of the two pathways for the mechanism students were expected to describe is presented in Figure 8.2. As students were given starting materials and products, the learning objective for the mechanistic description was for students to demonstrate their reasoning for the reaction mechanism. We limited the focus of this study to students' descriptions of the amide hydrolysis mechanism.

- Describes history of thalidomide used to treat morning sickness in pregnant women
- Identifies present value of using thalidomide to treat cancer and leprosy
- Identifies the acid hydrolysis mechanism to produce two products
- Specifies student's role as an organic chemist seeking to identify an analog of thalidomide that prevents hydrolysis
- Provides the writing goal to produce a description explaining the structure and reactivity of thalidomide toward hydrolysis

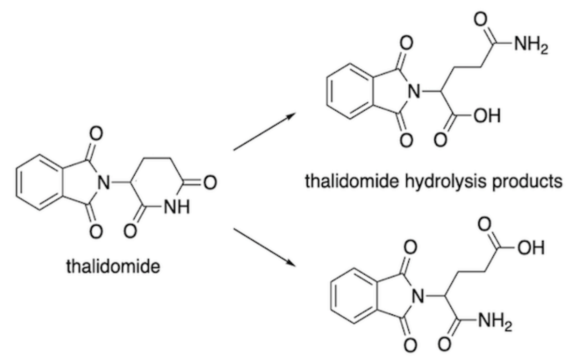


Figure 8.1 Relevant prompt components and the starting material and products for the reaction students were asked to describe and explain.

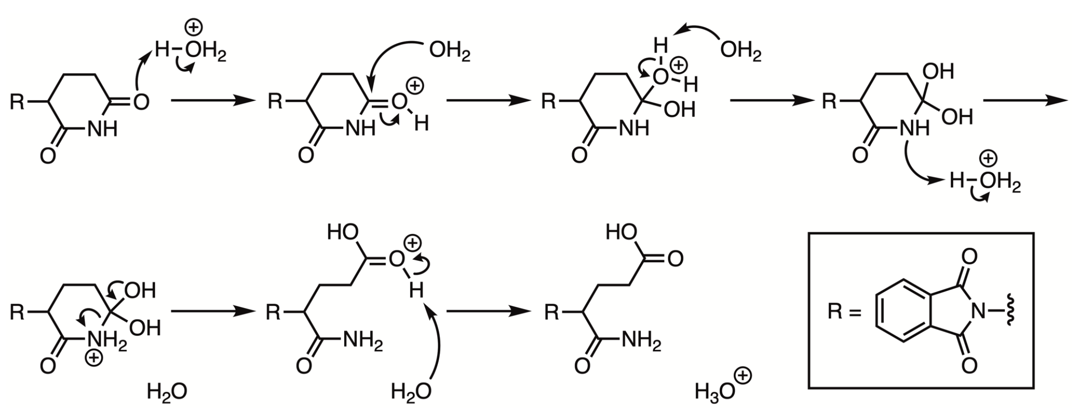


Figure 8.2 The acid-catalyzed hydrolysis of one of the thalidomide molecule's amide carbonyls. This is one of the mechanistic pathways students were expected to describe; the other pathway is the hydrolysis of the other amide carbonyl.

8.6.3 Writing-to-learn implementation

Students' first drafts were due on a Friday, after which students were randomly assigned to read and provide feedback for three of their peers in a double-blind peer-review by the following Monday. After receiving feedback, students were required to revise and resubmit the assignment by the end of the week. Students were able to ask questions and receive guidance on the assignment from the course writing fellows who were undergraduate students that had previously been successful in the course and were trained to provide feedback on content and writing. Grades for this assignment were determined independently of the present analysis.

8.6.4 Data collection

The data collected and analyzed from the WTL assignment were students' final drafts. Before collecting any data, the Institutional Review Board granted approval for the study and the participating students provided consent. Students' final drafts were the only data source included because students' revised writing best captures the features they found important to include in their mechanistic descriptions after receiving peer feedback and revising their work. Analyzing only the final drafts was done to focus on the writing that best represented students' knowledge after engaging with the cognitive processes of writing as facilitated by the structured peer-review process.

8.6.5 Data analysis

Analytical framework. We conducted the writing analysis by coding students' final revised drafts from the WTL process. Analysis was guided by an analytical framework presented by Russ et al.,[51] originally adapted from Machamer, Darden, and Craver's generalized description of a mechanism.[52] The framework provides a coding scheme for discourse analysis to identify the presence of mechanistic reasoning. The coding scheme is in the form of a logical hierarchy of codes for features expected to be present in a mechanistic description. This analytical framework was chosen for its focus on identifying features necessary for mechanistic reasoning in students' discourse, and because it aligned with the prompt in which students were asked to explain the acid hydrolysis mechanism.[51]

This framework was successfully used in other chemistry education research studies focused on mechanistic reasoning in the context of organic chemistry[13,32] and in the context of general chemistry.[21] Caspari et al.[32] utilized the framework to analyze organic chemistry students' ability to propose mechanisms while Caspari et al.[13] similarly used the framework to analyze students' construction of accounts relating structural changes to reaction energies, both in interview settings. Moreira et al. utilized the framework to analyze high school students' written responses after being given ten minutes to respond to a brief writing assignment eliciting mechanistic explanations of freezing point depression.[21] The present study similarly adapts this framework for recognizing students' mechanistic reasoning, but differs in that it is focused on written descriptions of the amide acid hydrolysis reaction mechanism. The adaptation of this framework to organic chemistry students' writing about more complex reaction mechanisms is valuable for understanding how these students think about and understand chemistry principles as applied to organic reactions. Furthermore, this study is differentiated by the WTL process used to promote students' engagement with the cognitive processes of writing.

The framework presented by Russ et al. is centered around entities and activities.[51] Entities are defined as the things which are involved in a mechanism.[51,52] In terms of organic reaction mechanisms, entities are electrons, atoms, and molecules.[13] Activities are defined as the actions entities take to produce change.[51,52] For organic reaction mechanisms, activities include the movement of electrons and the breaking and forming of bonds that produces structural change over the course of the mechanism.[32] The original framework described by Russ et al. included seven hierarchical levels—(1) describing the target phenomenon, (2) identifying setup conditions, (3) identifying entities, (4) identifying activities, (5) identifying properties of entities, (6) identifying organization of entities, and (7) chaining.[51]

The coding scheme adapted from this framework, located in 8.11.2 Appendix 2, Table 8.1 and detailed in the results and discussion, was developed by deductively coding for features expected in students' writing for each level of the hierarchy and open coding for additional features present in students' writing. Early in the coding process, the authors decided to code on a sentence-level grain size with the allowance that all appropriate codes would be applied to each sentence. This grain-size was chosen so we would be able to analyze what features were present, how frequently they appeared, and how often they co-occurred with other features. The coding frame began with the first sentence in a students' response in which a code could be applied and ended when the response shifted to answering another part of the prompt.

We conducted the initial coding (which included deductive and open coding in tandem) on a randomly selected subset of student responses, using constant comparative analysis to ensure all features were represented in the coding scheme and to clarify coding definitions.[53,54] The first and second authors worked in conjunction to develop the coding definitions, and other members of the research team with knowledge of mechanistic reasoning in organic chemistry assisted with further refinements. Improvements made to the coding scheme included incorporating codes developed from the open coding into the appropriate level of the hierarchical coding scheme. For example, in our deductive coding we did not include students' descriptions of the connectivity of starting materials and reaction intermediates, but it was a feature present in many responses. Thus, this feature of students' writing was included in the open coding and later integrated into the identifying setup conditions category of the hierarchical coding scheme. The choice was made to expand what was included within the setup conditions category beyond what was expected, as descriptions of connectivity relate the organization of atoms bonded together. This aligns with the setup conditions category, as specific connectivity is a requirement for particular mechanistic steps to occur. Furthermore, the way students wrote about and described connectivity during the course of the mechanism aligned with this category of the coding scheme, as their descriptions for products of one mechanistic step operationally served as the setup conditions for the next mechanistic step in the reaction. We combined and reorganized other codes from the deductive and open coding into the adapted coding scheme in a similar fashion. Additionally, we determined that some aspects of the original framework were not appearing in students' writing at the sentence level and thus we did not incorporate these into the coding scheme. The process of developing the coding scheme continued until saturation was reached.[55] In total, we coded 163 responses, representing 30% of the entire dataset.

The finalized coding scheme included four broad categories corresponding to four levels of the original framework that reflect the features necessary for engaging in mechanistic reasoning: (1) describing the target phenomenon, (2) identifying setup conditions, (3) identifying activities, and (4) identifying properties of entities. Codes relating to general descriptions of hydrolysis or the two reaction pathways leading to the two hydrolysis products were placed in the category of describing the target phenomenon. The identifying setup conditions category included codes relating to specifying the reaction medium or describing the structure or connectivity of starting materials, intermediates, and products. The third category, identifying activities, included codes relating to descriptions of electron movement or descriptions of bonds being broken or formed. The final category included the properties of entities—such as being acidic or basic, nucleophilic or electrophilic, or formally charged—that students identified in their mechanistic explanations. To illustrate the application of the coding scheme, two example student responses, with the applied codes indicated, are provided in 8.11.3 Appendix 3, Figure 8.14.

We did not include the third level of the original hierarchy, identifying entities, in the adapted coding scheme because the relevant entities (electrons, atoms, and molecules) were inherently coded for in other categories of the coding scheme. In other words, students never simply identified the entities without also describing their properties or the activities in which they were engaged. We also did not include the final two levels of the original framework—identifying organization of entities and chaining. Identifying the organization of entities was not included because of the category's focus exclusively on the spatial organization of entities as they are interacting during a mechanistic step, a feature which did not present itself in the students' writing. It is possible that whether or not students attend to the organization of entities depends on the mechanism—for instance, it might be present in mechanisms where there is a difference in stereochemical outcome depending upon the spatial organization of molecules as they interact (e.g., a unimolecular elimination reaction), or where spatial orientation during a mechanistic step might be described (e.g., the backside attack during a bimolecular substitution reaction). Chaining, defined as an explanation of how each mechanistic step leads to the next or why steps occur in the order that they do,[51] did not appear distinctly in student responses aside from the ordering of mechanistic steps. There was little variety in the ordering of mechanistic steps in students' writing, and analyzing chaining was not an insightful avenue of analysis in the present study due to this uniformity. It is likely that chaining pertains primarily to non-written descriptions of mechanisms in which students are proposing unknown mechanisms, or to written descriptions when students do not have the opportunity to refer to outside resources or revise their assignments after peer-review. Notably, chaining was the focus of the coding scheme presented by Caspari et al.,[32] in which students were proposing familiar and unfamiliar mechanisms during an interview. It is also possible that chaining was not identified due to the sentence-level grain size for coding, as chaining requires recognizing connections between mechanistic steps that might only be apparent across multiple sentences. Though chaining was likely present in students' thought processes regarding the hydrolysis mechanism, it was not necessarily identifiable in the conducted analysis.

Reliability. After finalizing the coding scheme, two authors independently coded 50 randomly selected responses to assess inter-rater reliability. The two coders met to check agreement, discuss codes, and make minor changes to the coding definitions to ensure the application of the coding scheme was clear. The fuzzy kappa statistic, a modified version of Cohen's kappa that allows for individual coding units to have multiple codes applied, was used to measure the reliability of the coding scheme.[56] For the 50 responses coded by two authors (representing 30% of the coded data), the fuzzy kappa statistic was 0.81, indicating near perfect agreement.[57]

Post-coding analysis. After coding students' writing and assessing reliability, we performed further data analyses with NVivo 12 and RStudio to understand the results of the coding.[58,59] First, we examined the total number of responses for which each code was applied at least once to determine how many students were incorporating each code. We additionally examined the frequency data relating how often each code was applied to each response. For this data, we calculated descriptive statistics across the set of responses in which the code appeared to characterize the general trends for how many sentences reflected each code within a response. We also calculated descriptive statistics for response length (in sentences) and total number of codes applied to each response.

Lastly, we examined the co-occurrences of codes to develop a more detailed understanding of how students were reasoning through the acid hydrolysis mechanism. To do this, we calculated a metric called lift, an association rule which measures the degree of dependence between two items, for each pair of codes. These values are useful to determine which pairs of codes were appearing together more or less than probabilistically expected. Lift is defined as

$$\frac{P(\mathrm{A},\mathrm{B})}{P(\mathrm{A})\cdot P(\mathrm{B})}$$

where $P(\mathrm{A})$ is the probability of code A appearing, $P(\mathrm{B})$ is the probability of code B appearing, and $P(\mathrm{A}, \mathrm{B})$ is the probability of code A and code B appearing together.[60] We extracted the frequencies of each code and the frequencies of co-occurrence for each pair of codes from the coding results. Then, as the sentence was the grain size for coding, we determined probabilities by dividing the appropriate frequencies by the total number of sentences coded. We then used the probabilities to calculate lift, which compares the observed probability of two codes appearing together, $P(\mathrm{A}, \mathrm{B})$, to the expected probability of two codes appearing together, $P(\mathrm{A}) \cdot P(\mathrm{B})$. Hence, lift measures whether codes appear together more or less than probabilistically expected. Lift values are interpreted by whether they are greater than, less than, or equal to 1.0. Lift values greater than 1.0 indicate that codes appear together more often than expected (e.g., lift of 2.0 indicates that the codes appear together twice as often than they would due to chance), while lift values less than 1.0 indicate that codes appear together less often than expected (e.g., lift of 0.2 means the codes appear together one-fifth as often as they would due to chance). A lift of 1.0 indicates the two codes in question appear together as often as expected due to chance (i.e., that they are independent of one another).

8.7 Results and discussion

The results from analyzing students' written descriptions of the hydrolysis reaction are drawn from the application of the coding scheme adapted from Russ et al.,[51] specifically by examining the prevalence and co-occurrences of codes within students' responses. The codebook is structured with four broad categories, each containing codes that indicate the specific features of students' writing corresponding to each category. These categories relate to the different components necessary for mechanistic reasoning present across the set of responses. We first report the percentages of responses in which each of the broad categories appears. Next we provide a detailed description of each category, focusing on the codes used to support claims made throughout the section. Lastly, we include an analysis of the co-occurrences of codes to make inferences about students' mechanistic reasoning for the acid hydrolysis mechanism.

8.7.1 What features are present in students' written mechanistic descriptions?

To examine the features present in students' written descriptions, we first observed how often each of the four broad categories of the coding scheme appeared in responses across the dataset. For these categories, 99% of responses included at least one description of the target phenomenon, 96% included an indication of setup conditions for the mechanism, 100% included a description of an activity taking place over the course of the mechanism, and 95% included an identification of the properties of entities. The high percentages of students incorporating each of these components necessary for mechanistic reasoning in their response indicates that the assignment, in general, successfully elicited descriptions of the acid hydrolysis mechanism. Since the majority of these features were present across responses, these values also suggest that the majority of students likely engaged in some form of mechanistic reasoning, which was the objective of the WTL assignment.

8.7.2 How do students write about the features present in their mechanistic descriptions?

Next we describe and provide examples of codes to illustrate how students appealed to each category of a mechanistic description. The reported percentages indicate the proportion of students including particular features in their response at least once. The full coding scheme, with definitions and examples for every code, can be found in 8.11.2 Appendix 2.

1. Describing the target phenomenon. The category of describing the target phenomenon included two codes, identified in Figure 8.3. Nearly all students included some description of the target phenomenon, and 98% included an overview of the reaction. Students' writing that contained an overview of hydrolysis included simply naming the reaction about to be described or identifying the two hydrolysis products. Some students also included a general description of hydrolysis, such as "Hydrolysis is the breakdown of a compound which proceeds as a result of water reacting with a carbonyl group."

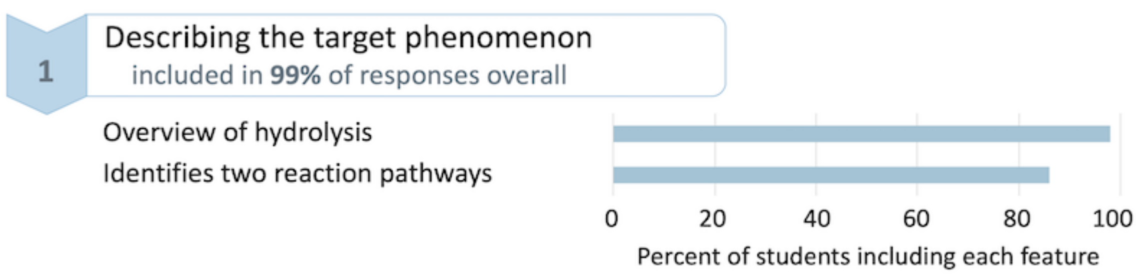


Figure 8.3 Percent of students incorporating features that describe the target phenomenon.

Students identified the two reaction pathways by stating an explanation, however minimal, of why two products were formed—such as "Two different hydrolysis products can be made based on which carbonyl gets attacked, but the mechanism is the same." Note that this example was also coded with providing an overview of hydrolysis, as it also states that there are two hydrolysis products. Students' responses might also have included language suggestive of the existence of multiple reaction pathways without explicitly making the connection to the two hydrolysis products, as in statements such as "This hydrolysis reaction can occur with either one of the carbonyl groups present on the ring." Notably, 14% of students did not make reference to the two reaction pathways leading to the different hydrolysis products identified in the writing assignment. This suggests that some students are not considering or placing enough importance on alternative, essentially equivalent, reaction pathways even when the results of these pathways are presented to them.

2. Identifying setup conditions. The level for identifying setup conditions included codes that pertained to the reaction medium or the connectivity of the molecules involved in the mechanism, as specified in Figure 8.4.

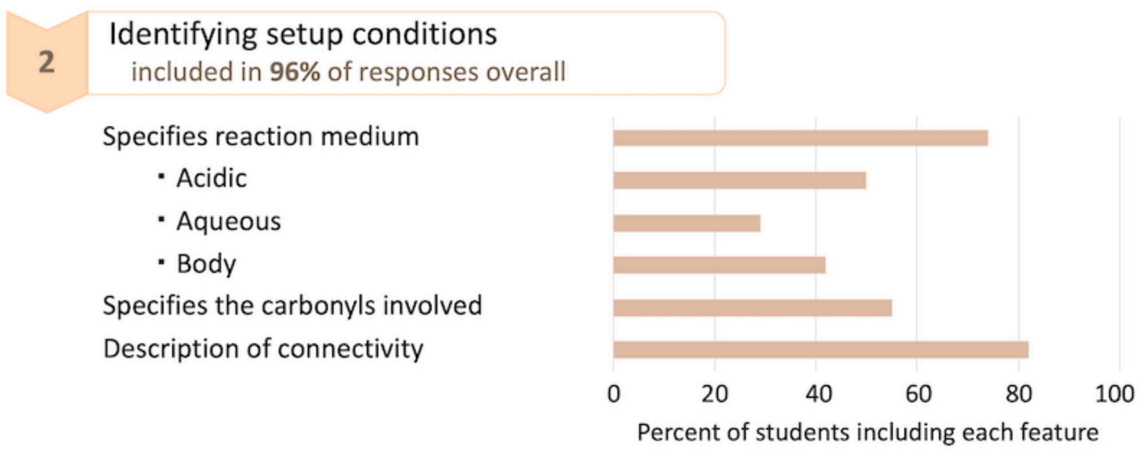


Figure 8.4 Percent of students incorporating features that identify the setup conditions.

Students described the acidic reaction medium by including phrases such as the "acid present in solution," the "acidic environment" or the "acidic conditions." Students similarly described the aqueous reaction conditions. As shown in Figure 8.4, 74% of responses incorporated at least one of the codes relating to the reaction conditions—and of that 74%, only 50% identified the reaction as occurring in acidic conditions and only 29% identified the reaction as occurring in aqueous conditions. From these percentages, it is clear that not all students are recognizing the value of identifying the reaction conditions in their mechanistic descriptions despite the importance of reaction conditions for understanding a mechanism.

Students specified the carbonyls involved by identifying the location on thalidomide where the hydrolysis reaction was taking place. They did this by providing some spatial description to identify which of the four carbonyls was reacting, such as "carbonyl in the 6-membered ring" or "carbonyl that is closest to the stereocenter" or "furthest away from the aromatic ring." This code only appeared in 55% of responses, suggesting that nearly half of the students did not pay sufficient attention to differentiating the reactive and non-reactive carbonyls.

Many students provided a description of the connectivity for the starting materials, intermediates, or products of the reaction. Descriptions of connectivity ranged from being relatively detailed (e.g. "the nitrogen atom that is part of the imide group is attached to a hydrogen atom") to including only reference to a functional group (e.g., "the Thalidomide molecule has two amide groups" or "…creating a hydroxyl group"). Students also included more general descriptions of connectivity such as "At this moment, we have a neutral tetrahedral intermediate." Descriptions of connectivity for the starting materials and intermediates are considered setup conditions for the mechanism, as such descriptions help the reader identify the connectivity required for each step of the mechanism to take place.

3. Identifying activities. The level for identifying activities included codes for descriptions of electron movement and changes in bonding. As seen in Figure 8.5, 99% of responses included some description of electron movement, while 100% of responses included some description of changes in bonding.

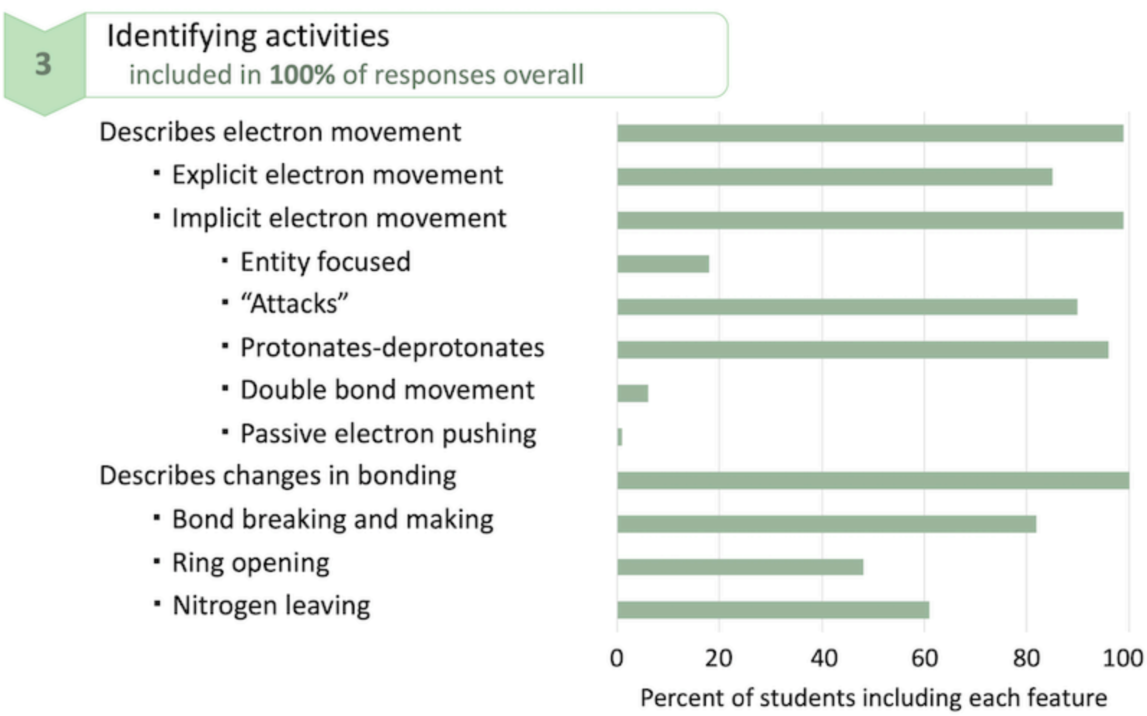


Figure 8.5 Percent of students incorporating features that serve to identify activities.

Students described electron movement both explicitly and implicitly. Explicit descriptions included students' reference to "electrons" or "lone pairs" when describing the movement of electrons. Implicit descriptions were those which did not explicitly refer to electrons, and were subdivided into codes for descriptions (a) focusing on the entity, (b) using variations of the word "attacks," (c) using variations of the words "protonates" and "deprotonates," (d) suggesting the movement of a double bond, and (e) mentioning passive electron pushing. Students' descriptions of entity-focused implicit electron movement included instances when the subject of a sentence describing a mechanistic step was something other than electrons (e.g., "One of the hydroxyl substituents forms a double bond. . ."). Students' use of the word "attacks" is a special case of this code in which the subject of the sentence was something other than electrons and the verb of the sentence was "attacks" (e.g., "Water then attacks…"). Students also described mechanistic steps using variations of the words "protonates" or "deprotonates." Descriptions indicating the movement of double bonds were those which described the movement of a pi bond rather than the movement of electrons in a pi bond. The code for electron pushing was applied when students passively described electron movement, in the sense of identifying something other than the entity involved in the mechanism performing the action (e.g., "The oxygen in the water molecule then attacks the carbon in the carbonyl, which, through electron pushing, forms a tetrahedral intermediate…"). Despite its infrequent appearance, this code remained in the codebook because it was an artefact of students' language use aligning with prior findings in the literature which suggest that students find the electron pushing formalism to be simply an academic exercise with little physical meaning.[5,6] It is promising that the potentially more problematic codes for descriptions of implicit electron movement appeared infrequently.

Explicit descriptions of electron movement were present in 85% of responses, while at least one of the codes for implicit descriptions of electron movement was present in 99% of responses. That a majority of students explicitly referred to electrons is a promising finding, indicating that the WTL assignment encouraged students to make connections between mechanistic steps and the movements of electrons. This suggests that, during the process of writing, students are attentive to the physical meaning of mechanistic steps, as opposed to prior studies that have shown students to not associate physical meaning when using the electron-pushing formalism.[5,6] However, 15% of students did not, in any sentence of their mechanistic description, identify the movement of electrons to describe a mechanistic step, while nearly every student included implicit descriptions of electron movement. Note that nothing is inherently wrong with implicit descriptions of electron movement; these descriptions simply do not indicate with certainty whether students are conceptualizing mechanistic steps as occurring due to the movement of electrons. It is notable that the most common codes for implicit electron movement are those for using variations of the words "attacks," "protonates," and "deprotonates," as practicing chemists and instructors frequently use these words when describing mechanisms. This provides evidence that students are using appropriate language when describing mechanistic steps.

The other set of codes categorized as identifying activities included descriptions of changes in bonding, as indicated in Figure 8.5. Students commonly did this using phrases such as "the bond between the nitrogen and carbon breaks" or "A lone pair from the oxygen reforms the carbonyl double bond." These descriptions can be thought of as a counterpoint to the aforementioned code for descriptions of connectivity in that this code was applied to active descriptions of changes in connectivity while the other code was applied to descriptions of connectivity before or after mechanistic steps. Students largely included descriptions of bonds being broken or formed, but 18% of responses contained no explicit description of this. Many students also referred to surface features of molecules to describe changes in bonding for the ring-opening step, with 48% of responses describing changes in bonding as a ring opening and 61% of responses describing changes in bonding as the nitrogen leaving. It is not necessarily incorrect to describe changes in bonding in terms of these surface features; however, it does suggest that some students may be overlooking the fundamental changes occurring in mechanisms—the bonds being broken and formed—in favor of paying attention to the more obvious surface features (such as the ring opening or nitrogen leaving, changes in bonding which result in obvious structural change).

4. Identifying properties of entities. The final level of the coding scheme, shown in Figure 8.6, included codes that identified the properties of the involved molecules that students used in their explanation of the acid hydrolysis mechanism. Students identified acids and bases by explicitly identifying the entity performing an activity as an acid or base or by referring to a mechanistic step as an acid–base reaction. Students identifying nucleophilicity or electrophilicity included specific reference to the molecules involved in a mechanistic step acting as either nucleophiles or electrophiles, occasionally including definitions of these words as well. Students identified charges by using words such as "positive," "negative," or "neutral" to describe a molecule acting in the mechanism. Some students included slightly more detailed explanations of charge, such as “The positive oxygen activates the carbonyl making the carbon a partial positive.”

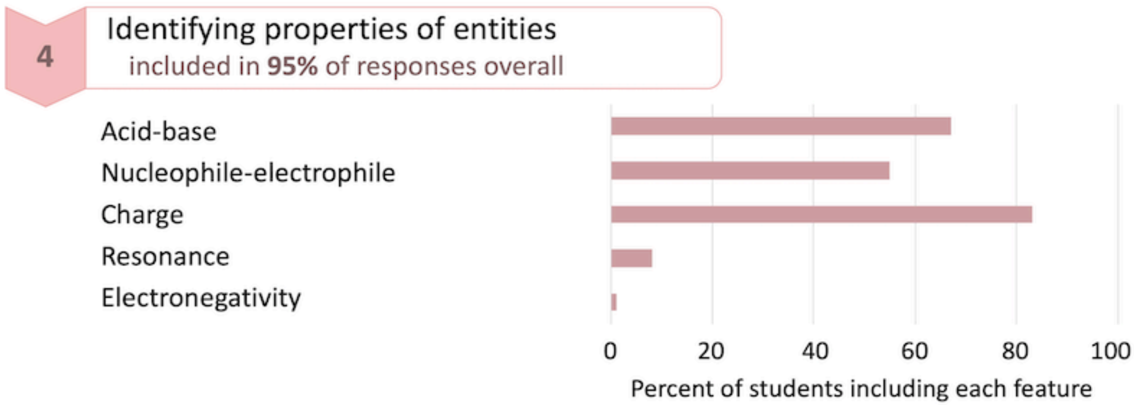


Figure 8.6 Percent of students incorporating features that appeal to chemical concepts.

As illustrated in Figure 8.6, only 55% of responses appealed to the properties of reacting molecules as nucleophiles or electrophiles, which is a fundamental property for explaining an acyl transfer mechanism. Instead, more students (67%) appealed to the properties of molecules as acids or bases. This is not surprising, as many of the reaction steps are protonations and deprotonations. Furthermore, acid–base chemistry is a topic that is introduced in general chemistry, so students in organic chemistry are likely more familiar with thinking of molecules in terms of acids and bases than in terms of nucleophiles and electrophiles. An even higher percentage of students (83%) appealed to the charged nature of reacting species. Again, this is not surprising since charges are explicit, surface features of molecules that change during the mechanism and are perhaps the simplest way for students to connect the movement of electrons to the properties of molecules. The relative percentages of students appealing to these three different properties of molecules aligns with prior studies in which students were found to rely on charges when considering mechanisms.[9,13,18,61]

The remaining codes in the category—identifying resonance or electronegativity—appeared less frequently. Students identified resonance by applying the concept either correctly (e.g., “The positive charge on the oxygen atom is stabilized through resonance”), somewhat correctly (e.g., “The resonance form of this molecule results in a positive charge…”), or incorrectly (e.g., “The electrons from the double bond resonate onto the oxygen”). Some responses also appealed to the electronegativity of atoms to describe electron density. It is somewhat surprising that few students identified resonance or electronegativity, as prior studies have shown that students often use these concepts to guide their mechanistic thinking.[6] However, it is unclear whether this is due to the specific mechanism students described or the nature of producing a written mechanism.

Overall, the results for the first two research questions (summarized by the complete coding scheme in 8.11.2 Appendix 2 and the appearance and frequency data in 8.11.4 Appendix 4, Table 8.2) indicate that while most students are including the components necessary for mechanistic reasoning as identified in the adapted coding scheme, there is considerable variety in how students include each of these components. Furthermore, despite promisingly high percentages of students appealing to each level of the coding scheme, the results draw attention to the codes within each category for which fewer students are incorporating particular components necessary for mechanistic reasoning.

8.7.3 What inferences about students' mechanistic reasoning can be made by analyzing co-occurrences of the features necessary for mechanistic reasoning?

In addition to what features were present in students' responses and how frequently these features appeared, we also examined the frequencies in which codes co-occurred with one another. We did this to make inferences about how students were engaging in mechanistic reasoning in their written explanations of the acid hydrolysis mechanism, specifically by examining how students combined properties of entities with the activities during the mechanism. In order to assess which pairs of codes were co-occurring in a meaningful way, we calculated the lift for each pair as described in the methods. The lift values and co-occurrence frequency data for all pairs of codes are presented in 8.11.5 Appendix 5, Figure 8.15 and Figure 8.16. From examination of the co-occurrence data, particular themes arose that are each supported by specific lift values and sets of Venn diagrams. Each of these themes are described below.

1. Students' writing provides empirical evidence for the hierarchical nature of the framework for identifying components necessary for mechanistic reasoning. The hierarchical nature of the analytical framework follows directly from the hierarchy of codes originally described by Russ et al..[51] Furthermore, this hierarchical relationship is implied by prior studies of students' reasoning abilities that progress from descriptive to relational to linear causal to multicomponent reasoning.[10,31,37,38] These studies are aligned with research conducted by Moreira et al. in which the hierarchical relationships between features of a mechanistic description were present in their classification of students' reasoning from "descriptive" to "emerging mechanistic."[21] In this study, the components increasingly built upon one another and connected to each other as the sophistication in students' reasoning increased.[21] Our results corroborate these prior studies by providing further empirical evidence of the hierarchical nature of the components necessary for mechanistic reasoning. Specifically, the lift values calculated between codes within the same category and between codes within neighboring categories identify that such pairings generally co-occur more frequently than pairings from non-neighboring categories. Overlaps within and between the first two categories of the coding scheme can be seen in Figure 8.7. The co-occurrences between these categories are evident with the high lift for providing an overview of hydrolysis with identifying two reaction pathways (1.57) and with the codes for specifying the reaction medium (ranging from 1.15 to 2.45). There are also high lift values between the codes for specifying the reaction medium (ranging from 2.94 to 3.42), showing the overlap between codes within the second category.

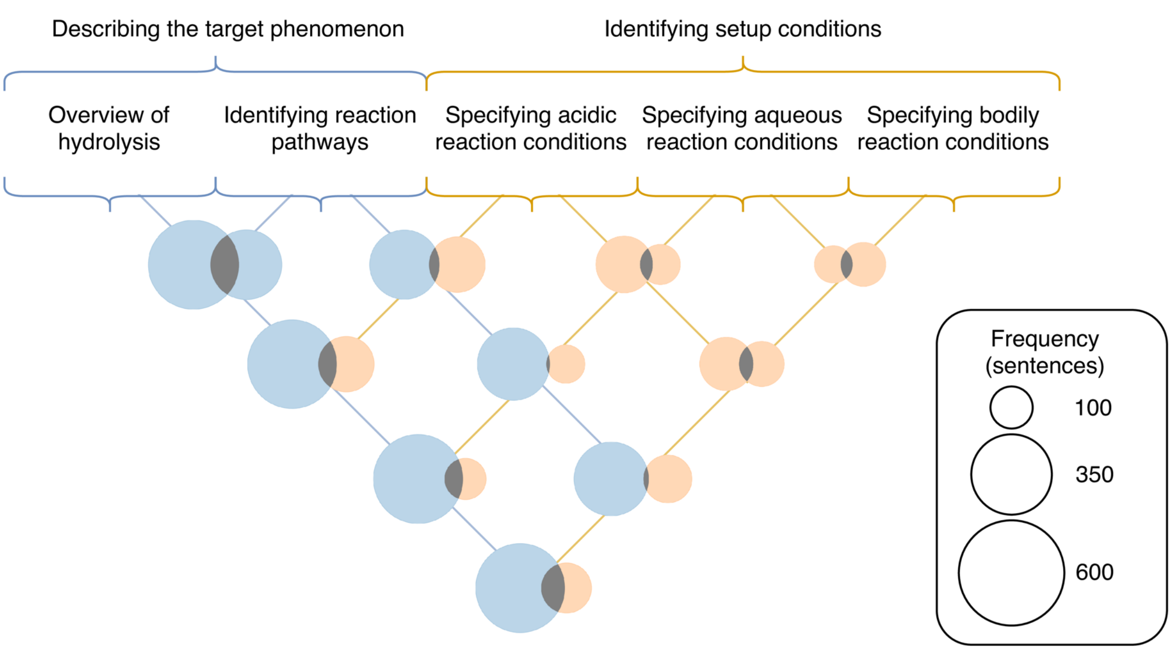


Figure 8.7 Venn diagrams between codes for describing the target phenomenon and identifying setup conditions. Overlaps indicate the number of sentences in which both codes in the pair appear together.

There are similar trends between codes in the third category of the coding scheme (describing activities), with some notable co-occurrences as illustrated in Figure 8.8. First, explicit descriptions of electron movement had high lift with the code for implicitly describing electron movement with the word "attacks" (1.75). This is an artefact of when students used the word "attacks" followed by an explicit depiction of electron movement—such as the case when a nucleophile attacks an electrophilic carbonyl followed by the movement of the pi electrons onto the carbonyl oxygen. Explicit descriptions of electron movement also had high lift with the three codes related to the formation or breaking of bonds (2.34, 2.85, and 3.24). This finding aligns with prior research that has found students to be able to describe changes in bonding using electron movement.[18] In contrast, the codes for implicit descriptions of electron movement—using the word "attacks," "protonates," or "deprotonates"—had lift values below 1.0 for the codes related to the formation of bonds. This suggests that students' writing does not reflect that bonds are formed or broken in the processes of nucleophilic attacks, protonations, or deprotonations. Unsurprisingly, there were high lift values (3.40, 3.03, and 4.27) between the three codes related to the forming and breaking of bonds, as students often explicitly described the fact that bonds were being broken or made in conjunction with describing the surface feature changes of the ring opening or nitrogen leaving.

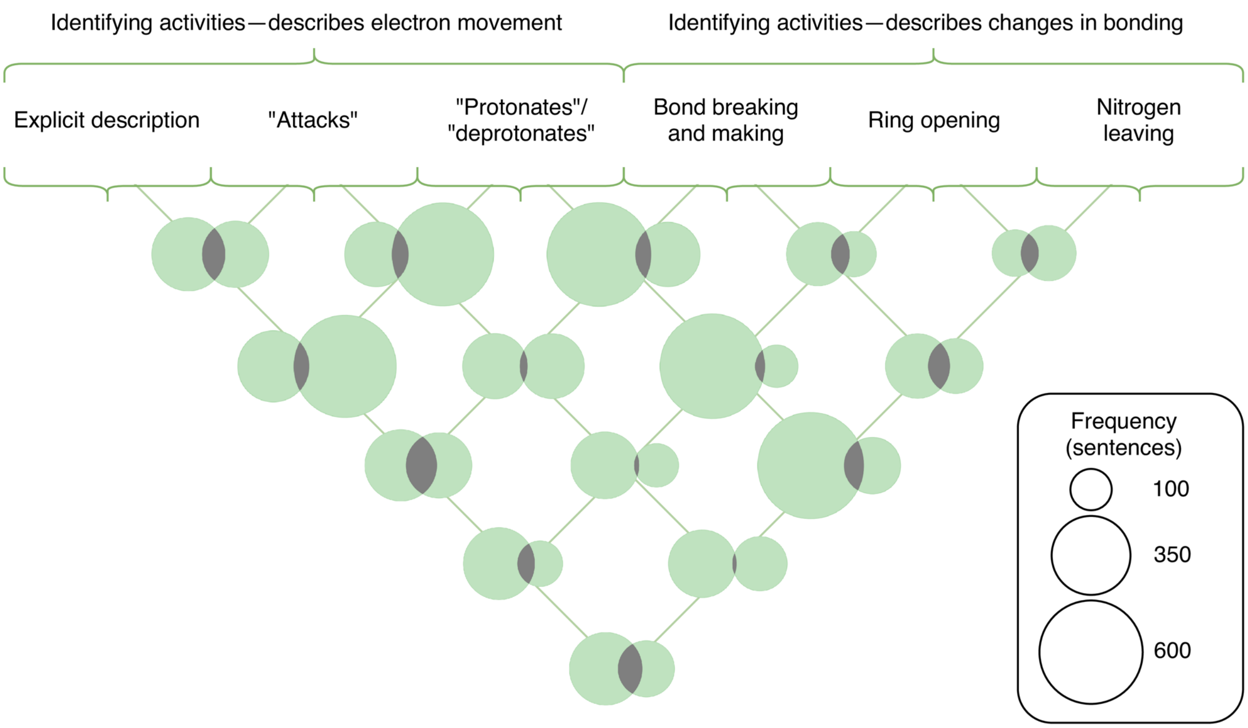


Figure 8.8 Venn diagrams between codes for identifying activities—split between the sub-codes for descriptions of electron movement and the sub-codes for descriptions of changes in bonding. Overlaps indicate the number of sentences in which both codes in the pair appear together.

Notably, the lift values were generally below 1.0 for codes in the first and second categories of the coding scheme paired with codes in the third and fourth categories. This result shows that the codes related to describing mechanistic activities (the third category) and identifying properties of entities (the fourth category) are largely independent of the codes for describing the target phenomenon (the first category) and identifying the setup conditions (the second category). The lift values below 1.0 provide further evidence for the hierarchical nature of students' mechanistic descriptions, as students included features from the first two categories alongside features from the last two categories less than expected by chance.

2. Students identified the two reaction pathways primarily by identifying divergence in the first step of the reaction. By examining the lift values between the codes identified in Figure 8.9, the connection students made between the reaction's first protonation step and the two reaction pathways was notable. The code for identifying reaction pathways had high lift (3.66) with only one code—the code for specifying the carbonyls involved in the reaction. The magnitude of the lift value suggests a strong dependence between these two codes, which is not surprising as the source of the two reaction pathways is directly connected to the two carbonyls present that undergo the same hydrolysis reaction. The co-occurrence between these two codes does, however, provide evidence that students are not merely stating that the reaction produces two products, but are connecting this outcome to the features of the starting material that are responsible for the two reaction pathways.

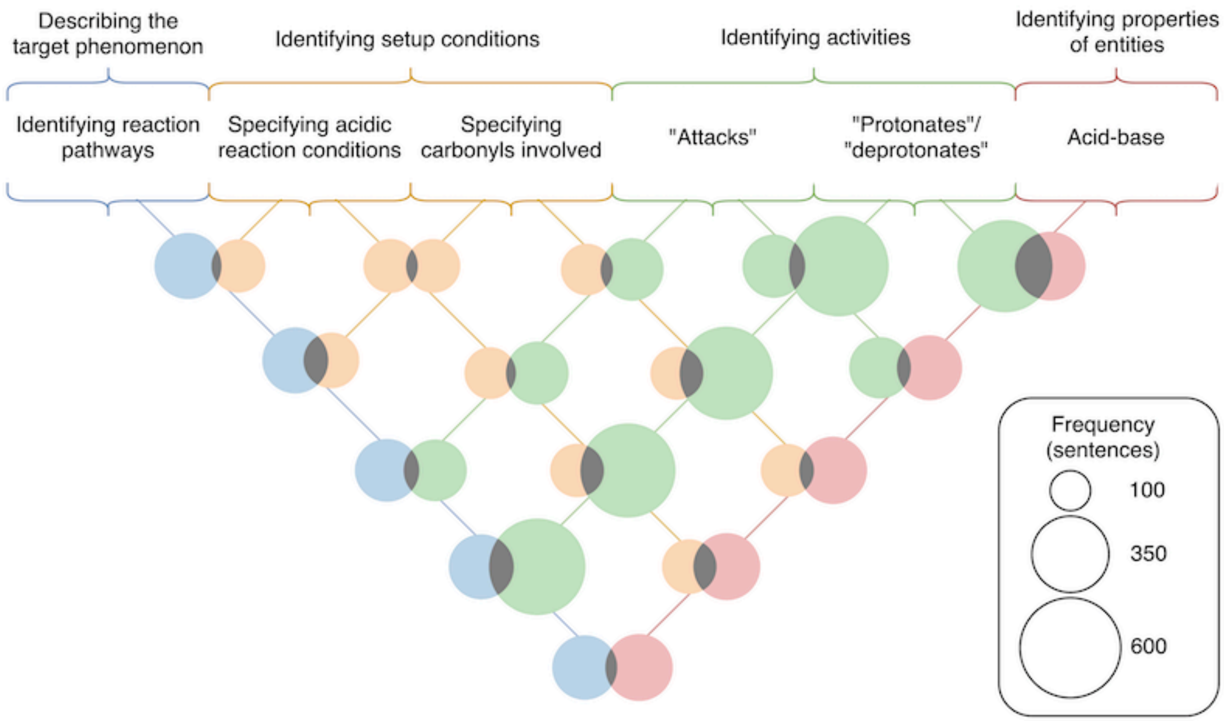


Figure 8.9 Venn diagrams between the codes relating to students' descriptions of the two reaction pathways yielding different hydrolysis products. Overlaps indicate the number of sentences in which both codes in the pair appear together.

The code for specifying the carbonyls involved in the reaction had high lift values with three other codes—identifying the acidic conditions (1.45), using the words "protonates" or "deprotonates" (1.54), and identifying entities as acids or bases (1.36). There were similarly high lift values between the other combinations of these codes (ranging from 1.47 to 2.15). The relationships between these codes show that students are making the logical connections between the acidic medium and the protonation steps in the mechanism—particularly the protonation of one of the two carbonyls that leads to one of the final products. This result differs from prior research by Caspari et al. and Petterson et al., in which students did not verbalize alternative mechanistic steps that lead to alternative reaction pathways.[12,32] This finding suggests that the WTL assignment, which included clear expectations to explain the formation of two products, elicited students' consideration of the alternative mechanistic pathways that they might not have considered otherwise.

Another observation is that the code for using the word "attacks" is relatively independent of the codes for identifying the reaction pathway or specifying the carbonyls involved (lift of 1.13 and 1.16, respectively). This independence is notable in light of the two ways students chose to identify the divergence in the reaction that leads to two products. The first, which the co-occurrence data suggests students did with more frequency, was to identify the divergence at the first step of the reaction—the protonation of one of the two carbonyls (e.g., "…the final product is determined by which oxygen is initially protonated" or "Depending on which amide is originally protonated, two hydrolysis products can form"). However, an alternative way that some students identified the divergence in the reaction was by considering which protonated carbonyl served as the electrophile in the nucleophilic attack by water (e.g., "The other hydrolysis product forms when water attacks the other carbonyl" or "The hydrolysis product depends on which carbonyl group on the 6-membered ring is attacked."). While the divergence at the protonation step is reflective of how this reaction mechanism might be drawn to show the formation of two products, the divergence at the step of nucleophilic attack suggests a potentially more nuanced understanding of the dynamic equilibrium between protonated and deprotonated species in acidic media, as the protonation step is likely to be more easily reversible than the nucleophilic attack. Hence, the lower co-occurrence between the codes for using the word "attacks" and identifying the two reaction pathways suggests that more students are writing the descriptions for alternative mechanisms as the individual mechanisms would be drawn, rather than locating within the description the most likely point of divergence. This result could indicate that some students do not have a full conceptual understanding of the dynamic nature of reactions, especially when reactions lead to similar products. The difference between these two descriptions could indicate differences in whether students perceive reactions to be occurring stepwise or in a more dynamic manner, a possibility that has emerged in other studies.[18]

Furthermore, the set of co-occurrences between identifying the acidic conditions, using "protonates" or "deprotonates," and identifying entities as acids or bases (with lift values ranging from 1.47 to 2.15) illustrates that students did make the connection between the acidic medium and the presence of a molecule acting as an acid to perform a protonation. This finding suggests that students engaged in reasoning that connected the acidic setup conditions to the molecules being in a protonated state through the mechanism of an acid–base reaction. Notably, there is no dependence between the acidic conditions code and the charge explanation code (lift of 1.06). This may be an artefact of students not making the conceptual connection between acidic environments and the presence of positively charged species. However, we might expect students to apply rule-based reasoning to directly make this connection using the rule that positive charges are associated with acidic reaction conditions, similar to students' rule-based-reasoning described in prior studies.[3,36,39] Hence, this result may suggest that the WTL assignment facilitated reasoning reflective of process-oriented rather than product-oriented problem-solving.

3. Students made appropriate connections between mechanistic steps and properties of entities. Another finding from examining the co-occurrence data is how students' descriptions of changes during a mechanism relate to the identified properties of entities involved in the change. These co-occurrences are illustrated in Figure 8.10. First, there is a large lift (4.14) between the code for using the word "attacks" and identifying entities as nucleophiles or electrophiles, meaning these two codes appeared together approximately four times more than expected by chance. There is also a demonstrated dependence between using the words "protonates" or "deprotonates" and identifying acids and bases (lift of 2.15) or charge (lift of 1.49). These are expected overlaps, as reactions between nucleophiles and electrophiles are typically described as the nucleophile "attacking" the electrophile and protonations and deprotonations are acid–base reactions which result in changes in charge. However, it is possible that students might have described entities as nucleophiles simply due to the fact that they attack another entity, rather than inferring the nucleophilicity from electronic properties (i.e., a lone pair of electrons or a partial negative charge). Similarly, students might have recognized acids and bases simply from the fact that they are engaged in an acid–base reaction rather than inferring their acidic and basic properties from structural features. Nevertheless, these co-occurrences provide evidence that students are using appropriate language to discuss the chemical properties related to particular changes occurring during the mechanism. While there are expected overlaps between the codes for describing electron movement and identifying properties of entities, the lift values are near or below 1.0 between the three codes for describing changes in bonding and the three most prevalent codes for identifying properties of entities (charges, acid/base, or nucleophile/electrophile). This pattern shows that students were appealing to the properties of entities to justify electron movement but were rarely using the properties of entities to justify changes in bonding.

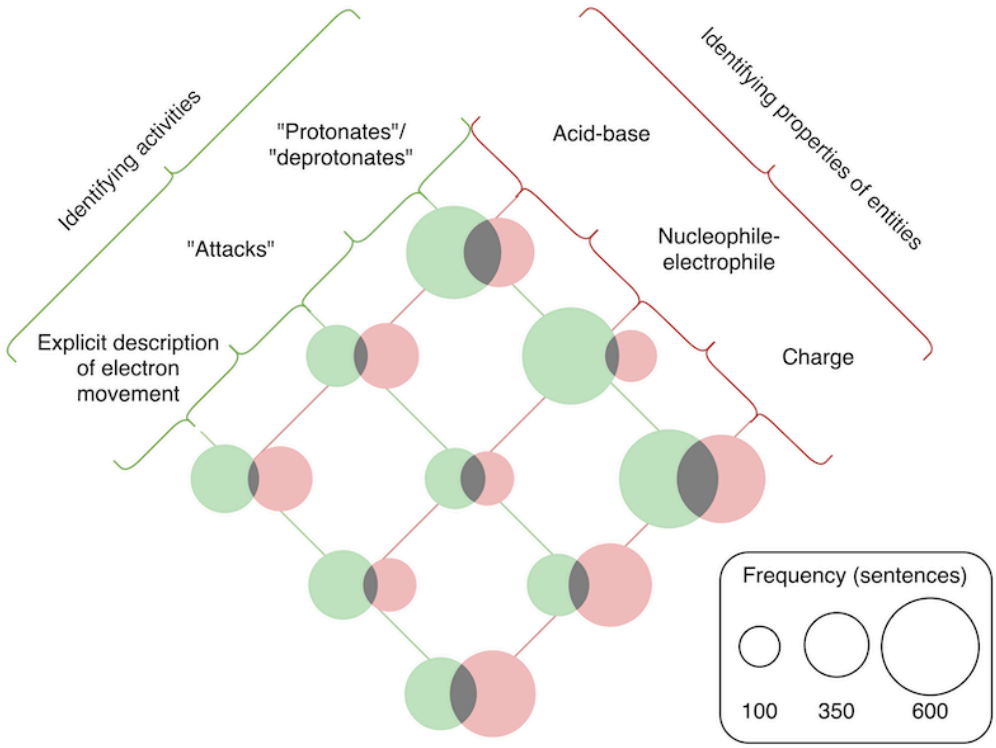


Figure 8.10 Venn diagrams illustrating the overlaps between codes for descriptions of electron movement and codes for identifying properties of entities. Overlaps indicate the number of sentences in which both codes in the pair appear together.

The lift values between different properties of entities and explicit descriptions of electron movement are also notable. While the lift values between explicit descriptions of electron movement and identifying nucleophiles/electrophiles or charges are slightly above 1.0 (1.19 and 1.32, respectively), the lift between explicit descriptions of electron movement and identifying acids/bases is below 1.0 (0.51). These values reveal a modest dependence between describing explicit electron movement and identifying entities by either their nucleophilicity/electrophilicity or charge. However, the overlap between explicit electron movement and identifying acids/bases is less than expected due to chance—meaning that when students identified acids/bases they were less likely to accompany that identification with explicit descriptions of electron movement (and vice versa). This finding suggests that students are appealing to Brønsted–Lowry acid–base theory more than they are appealing to Lewis acid–base theory, aligning with prior research regarding students' application of different acid–base theories.[12,28,33] The lack of appeal to Lewis acid–base theory is valuable to recognize in students' writing, as the Lewis theory is a concept necessary for mechanistic reasoning[29] and students who use Lewis acid–base theory are more successful at mechanism tasks.[10,62] In addition, the percent of overlap between explicit descriptions of electron movement and the identification of properties of entities is the largest for identifying charges. Together, these findings suggest that students are able to connect explicit—as opposed to implicit—descriptions of electron movement with more accessible or surface-level reasoning (identifying charges or using Brønsted–Lowry acid–base theory) as opposed to reasoning with more sophisticated concepts (identifying nucleophiles/electrophiles or using Lewis acid–base theory). Such a focus on surface features of reactants has been shown to engender rule- or case-based reasoning, and might be reflective of students' product-oriented approaches to problem-solving.[3,36,39]

Lastly, among the three most prevalent codes for the identifying properties of entities, the lift values are less than 1.0 for identifying nucleophilicity and electrophilicity in conjunction with both other commonly identified properties (acidic/basic and charge). The overlaps between these codes are presented in Figure 8.11. These co-occurrences indicate that identifying nucleophiles and electrophiles occurs most commonly with the absence of identifying other properties of entities, matching findings from prior research in which few students made connections between acids/bases and nucleophiles/electrophiles.[33] However, there is a high lift value (1.57) between identifying acids and bases and identifying charges, indicating that these constructs frequently occur together. This lift value provides further support for the hypothesis that students are more comfortable identifying the more familiar construct of charge or using Brønsted–Lowry acid–base theory—and even use them to complement each other. On the other hand, when students do identify nucleophiles and electrophiles, it is much less likely to be accompanied with identification of other properties of entities. This finding may reflect students' abilities to engage in integrated multicomponent reasoning only with certain properties of entities (i.e., being able to use charge and acid/base character simultaneously), but that these abilities are limited when considering properties such as nucleophilicity or electrophilicity.[31,37,38]

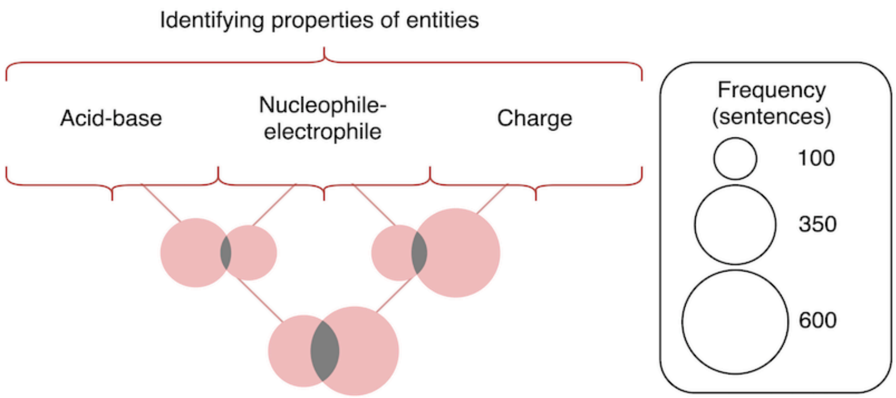


Figure 8.11 Venn diagrams between the codes for identifying properties of entities. Overlaps indicate the number of sentences in which both codes in the pair appear together.

8.8 Conclusions

We have described the analysis of student responses to a WTL assignment designed to elicit mechanistic descriptions of an acid hydrolysis reaction. Our study was guided by an analytical framework for discourse analysis grounded in the philosophy of science literature. Responses were coded for the presence of features necessary for mechanistic reasoning within the broad categories of describing the target phenomenon, specifying setup conditions, identifying activities, and identifying properties of entities. Our goal for coding was to provide a rich description of how students incorporated these features in their descriptions of the reaction mechanism. The second aspect of this research identified how these features co-occurred to make inferences about students' mechanistic reasoning. This analysis furthers our understanding of the way students think about reaction mechanisms in the context of a specific reaction. It has shown that, in general, the assignment successfully elicited complete mechanistic descriptions, as most students appealed to each level of components necessary for mechanistic reasoning as described by the coding scheme adapted from Russ et al.,[51] with 85% of students explicitly describing the movement of electrons. Additionally, trends in the co-occurrence data—in which codes within the same category or from neighboring categories generally co-occurred more often compared to codes from more separated categories—provided support for the hierarchical ordering of the components necessary for engaging in mechanistic reasoning.

A number of findings arose from analysis of the frequency and co-occurrence data presented which identify the features students did (or did not) engage with during the process of writing. First, there were notable percentages of responses that did not incorporate some of the important features of a description for the mechanism. Some students (26%) did not specify the reaction medium, indicating that these students are not recognizing the importance of the reaction conditions as they pertain to reaction mechanisms. Additionally, some students (14%) did not consider the two reaction pathways, even though the assignment explicitly requested an explanation for the formation of two products. For those students who did consider the two reaction pathways, there was evidence to suggest different interpretations of where the reaction diverged. Many students indicated the divergence at the first mechanistic step, while fewer students indicated the divergence at a later (more chemically reasonable) step, suggesting differences in students' understanding of the dynamic nature of reactions when considering multiple reaction pathways.

Perhaps most notable is that 45% of students made no reference to the reacting species as nucleophiles or electrophiles. In general, identifying charges was more prevalent than identifying properties of entities that allow for more sophisticated conceptual reasoning such as identification of nucleophiles and electrophiles or acids and bases. Furthermore, compared to other properties of entities, identifying nucleophilicity and electrophilicity occurred less often in conjunction with identifying other properties. The findings also showed that students more often made connections between charges and explicit descriptions of electron movement compared to other properties of entities. Explicit descriptions of electron movement were also frequently connected to descriptions of bonds being broken and formed, but this connection was not present for implicit descriptions of electron movement. In addition, when describing changes in the mechanism, identifying the properties of entities more frequently accompanied descriptions of electron movement than descriptions of changes in bonding. Another finding that presented itself throughout the data was that many students were using appropriate language to describe mechanistic steps. Students commonly used the word "attacks" when describing a nucleophilic attack and used variations of "protonates" or "deprotonates" in reference to acid–base reactions. This suggests that students were making appropriate connections between concepts across different categories of the coding scheme. Taken together, the findings from this research identify how students were engaging in mechanistic reasoning by revealing how students used or did not use different properties of entities in conjunction with descriptions of the activities and changes occurring over the course of the mechanism.

8.9 Limitations

This research is limited by a variety of factors. First, the generalizability of the results are limited by the context in which the research was conducted. Data was collected only from a single, selective institution. Students' mechanistic descriptions are likely influenced by their backgrounds, their instructors, and other factors which vary with institution. Specifically, the language used by instructors and the emphasis placed on particular aspects of mechanistic reasoning may influence students' written mechanistic descriptions.

The results are also limited by the data collected and the analytical framework. Since we only analyzed students' final drafts, the findings are limited to the evidence of students' reasoning demonstrated in their written work after the peer-review process. Some aspects of students' understanding may not be captured by examining their writing, and students' actual ability to reason through mechanisms could be greater or less than suggested by their writing. Also, the framework used to analyze students' writing did not assess the accuracy or correctness of the written mechanisms. Hence, the framework is limited to characterizing how students include the features necessary for mechanistic reasoning as opposed to whether or not their written mechanism is correct. The analysis is also limited in that no external measures of students' mechanistic reasoning were administered, so the research cannot suggest the efficacy of the WTL assignment to develop the capacity for reasoning.

Another limitation is that the framework was applied to a specific prompt eliciting students' mechanistic descriptions of a specific reaction mechanism. Descriptions of other reaction mechanisms might produce different results in terms of the prevalence of particular features; furthermore, writing to describe other reaction mechanisms might prompt students to incorporate additional features not included in the present analytical framework. Additionally, elements of prompt design likely influence the way students write about mechanisms. In particular, the features necessary for mechanistic reasoning not present in students' writing (e.g., identifying organization of entities) could be due to the specific mechanism or prompt examined in this study. The absence of these features could alternatively be an artefact of translating a mechanism into writing. This distinction is unclear and would require further research.

8.10 Implications

8.10.1 Implications for teaching

There are a number of implications for practice stemming from this work. First, this research presents a WTL assignment that successfully elicited detailed mechanistic descriptions, which, as suggested by the cognitive process theory of writing, can support students' learning. Additionally, the findings suggest that the language students use to write about mechanisms—and, tangentially, the way students think about mechanisms—is reasonably accurate and thus potentially influenced by the language instructors use when describing mechanisms. For example, students frequently used the word "attacks" to describe a nucleophilic attack, but it is not certain that students understand the implicit electron movement described when they write that a nucleophile "attacks" an electrophile. Therefore, it is important to be as explicit as possible that these words being used to describe mechanistic steps—words like "attacks" and "protonates"—are words that are implicitly describing the movement of electrons. Furthermore, it may be valuable for instructors to use words that more accurately represent molecular behavior—for example, replacing the word "attacks" with "collisions" when describing interactions between nucleophiles and electrophiles.

Building upon this observation, it is vital that instructors connect mechanistic steps to the underlying chemical properties driving mechanisms. The findings in this study suggest that students are able to say what is happening but not always able to explain why things are happening. This tendency suggests that instructors need to emphasize the appropriate use of fundamental chemistry concepts students should be thinking of when considering reaction mechanisms. In particular, instructors can place more focus on considering the nucleophilicity and electrophilicity of reacting species as a way to describe the flow of electrons in each step of a mechanism; this concept is perhaps the most fundamental way that practicing chemists think about mechanisms, but it was less common among students' written explanations in comparison to considerations of charges or acid–base chemistry.

In addition to carefully modelling for students all components of a mechanistic description when presenting a mechanism in class, further implications for practice could be to incorporate these components into mechanism questions on assignments or assessments. The four categories of features in students' mechanistic descriptions provide a natural scaffold for engaging students in mechanistic reasoning; these could be presented in the text accompanying a mechanism problem or could be made into problems themselves. For example, a problem asking students to provide a mechanism might include components where the student must identify the reaction conditions or describe the relevant properties of molecules driving particular mechanistic steps in addition to providing the electron-pushing diagram. Incorporating such questions into a problem will emphasize for students the components of a mechanism that practicing chemists are considering—the reaction medium, alternative reaction pathways, the properties of entities, etc.—as opposed to only emphasizing for students the electron-pushing formalism itself.

8.10.2 Implications for research

Prior research has identified differences in students' reasoning,[10,31,37,38] including identification of the hierarchical relationships between components of a mechanistic description.[21] The present research is the first study to use the lift metric to empirically demonstrate this hierarchical relationship between components. Furthermore, this study used lift to analyze a large set of written data to make inferences about students' mechanistic reasoning. This is valuable because it has allowed for the investigation of students' mechanistic reasoning at a larger scale, which in prior studies has been investigated using think-aloud interviews with limited numbers of participants. Generally, lift is a metric that can be applied in other settings to examine co-occurrences between codes in a qualitative coding scheme. It is applicable to any coding scheme in which multiple codes may be applied to a single unit of analysis and is valuable for identifying when code co-occurrences occur more or less than expected by chance. Hence, lift could be useful in analyzing coding results for any number of research studies utilizing a coding scheme.

Studies by Moon as well as Moreira examined students' writing to understand their reasoning[22] and mechanistic reasoning[21] in general chemistry and high school chemistry settings. This study expands on this work to examine students' responses to a WTL prompt eliciting explanations of an organic reaction mechanism. The methods presented in this study provide a route to access students' reasoning using qualitative methods to identify features in students' responses followed by a quantitative method to make inferences about their reasoning. This methodology could be used in similar studies of students' mechanistic reasoning to afford further insights. For instance, more specific coding of entities (e.g., specific functional groups) and their properties and activities could allow researchers to specifically characterize how students construct structure–property relationships. Such efforts could identify the sophistication of students' mechanistic reasoning by recognizing if students connect properties to function or simply associate specific structural features with particular mechanistic activities. This may be especially insightful in situations where students are proposing an unknown mechanism without access to outside resources, where they would be required to use these relationships to determine reaction progress. Furthermore, analyzing student writing, as opposed to their use of symbolic notation, could be applied to similar WTL activities engaging students in tasks of describing other organic reaction mechanisms. Doing so would broaden our understanding of how students reason through mechanisms and develop our understanding of the relationship between reaction type (e.g., hydrolysis *versus* substitution) and students' use of components necessary for engaging in mechanistic reasoning.

Additional studies are also needed to further explore the application of this framework in other contexts, with attention to variables such as institution, prompt design, instructors' use of language, and students' prior experience with organic chemistry. These variables, among others, may influence students' mechanistic descriptions. Beyond this, future research could include examining the effect of peer-review and revision on students' mechanistic descriptions by applying the framework to students' first and final drafts and examining changes in the presence of each feature of mechanistic reasoning. Another future direction could involve further examination of the data to identify if there are differences in mechanistic reasoning between students. For example, the features present in students' writing may correlate to their success in the course or relate to other factors linked to student performance. If this is the case, such writing assignments could be utilized as a tool for providing formative assessment to students in order to develop their mechanistic reasoning skills.

8.11 Appendices

8.11.1 Appendix 1. The writing-to-learn assignment

Thalidomide: a pharmaceutical Jekyll and Hyde

Thalidomide was widely used after World War II as a sedative and later as a treatment for morning sickness. Unfortunately, it was only after widespread use that it was discovered that thalidomide causes very serious side effects—in particular, birth defects such as phocomelia (limb malformation). The drug was banned in 1962 and these events resulted in important changes to the way the FDA approves drugs.

Despite the inherent dangers, thalidomide is now used for treatment of serious diseases, such as cancer and leprosy, when the benefit of treatment outweighs the inherent risks. It is now understood that thalidomide exists as two enantiomers; one is a teratogen and the other has therapeutic properties. Rapid racemization occurs at body pH and both enantiomers are formed at roughly an equal mixture in the blood, which means that even if only the useful isomer is used, both will form once introduced in the body. Furthermore, both enantiomers are subject to acid hydrolysis in the body and produce hydrolysis products that may or may not be teratogens depending on their structure. The structure of Thalidomide and two Thalidomide hydrolysis products are shown below in Figure 8.12.

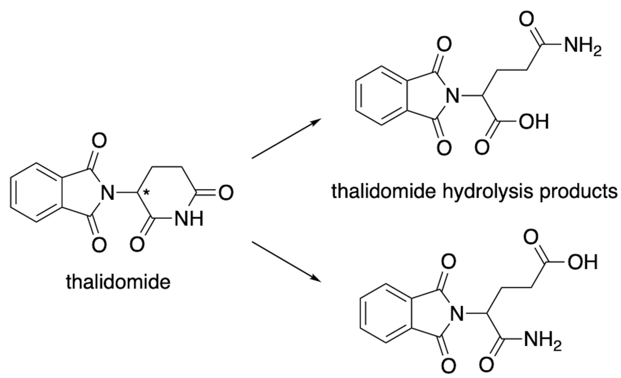


Figure 8.12 Thalidomide and thalidomide hydrolysis products. The stereocenter is shown (*).

You are an organic chemist collaborating with a team of other researchers from USC with the goal of testing Thalidomide analogs for cancer treatment. An analog is a compound that is very similar to the pharmaceutical target that has small structural differences. For example, *m*-cresol (shown in Figure 8.13 below) is an analog of phenol. Your goal will be to design a structural difference that will make the Thalidomide analog less reactive toward hydrolysis than Thalidomide. Your analogs will be tested for the inhibition of a pro-inflammatory protein mediator, which in elevated levels may be responsible for symptoms associated with the early stages of HIV.

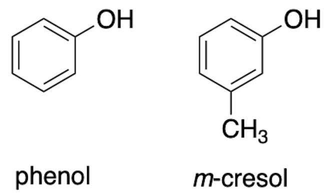


Figure 8.13 Example of an analog of phenol.

Although Thalidomide is warranted for treatment of some diseases, it would be preferable to identify an analog that has similar therapeutic qualities without the potentially devastating side effects. It is known that Thalidomide is easily hydrolyzed, and it has been proposed that one of the biologically active species may be one of the two possible hydrolysis products shown above. Thus it is important to propose analogs that are not readily hydrolyzed.

Your research team is drafting a grant proposal for the National Institute of Health. You must contribute a 500–750 word description explaining the structure and reactivity of thalidomide toward hydrolysis and the structural differences in proposed analogs that will make them inert to hydrolysis. The committee who will review the proposal is likely to be made up of scientists from disciplines including biology, chemistry and medicine. While they are experts in their own field, they may not be knowledgeable about organic chemistry, racemization, hydrolysis, or NMR spectroscopy.

When writing, you should consider the following:

1. Design one compound (thalidomide analog) that should be a pro-inflammatory protein mediator inhibitor. Explain.
2. Explain why it is important that thalidomide analogs do not have acidic protons at their stereocenters.
3. Explain the mechanism for acid hydrolysis of thalidomide to form the two hydrolysis products in Figure 8.12.
4. Describe how you would monitor hydrolysis of thalidomide by NMR.
5. Set the tone of your piece by placing your description in the context of the larger goal of developing a safer drug for the treatment of cancer patients.

6. You should consider carefully which organic chemistry terms you use and when you define or explain them. Remember, your collaborators are relying on you to clearly communicate your plan so that they can write a competitive proposal for funding from the NIH.

NOTE: you can choose to include drawings of either the mechanism or of your proposed analog. However, given your audience, your written explanation should be sufficient such that your proposed analog can be understood without the drawing. Your grade will be solely determined based on what you wrote.

8.11.2 Appendix 2. Coding scheme

Table 8.1 The finalized coding scheme used to analyze students' written descriptions of the hydrolysis mechanism

Category	Code name	Code name (shortened)	Definition	Exemplars
Describing the target phenomenon	Overview of hydrolysis	over	The sentence provides a broad description of the hydrolysis reaction.	"One reaction of thalidomide is an acid hydrolysis reaction" "Thalidomide is a compound which, when undergoing an acid hydrolysis reaction, can form two constitutionally isomeric products." "Hydrolysis is the breakdown of a compound which proceeds as a result of water reacting with a carbonyl group."
	Identifies two reaction pathways	idpath	The sentence identifies that the initial protonation and nucleophilic attack can occur at two carbonyls, which leads to two different products.	"Two different hydrolysis products can be made based on which carbonyl gets attacked, but the mechanism is the same." "The same general mechanism occurs when the other carbonyl is first protonated" "This hydrolysis reaction can occur with either one of the carbonyl groups present on the ring."
Identifying setup conditions	Specifies reaction medium—acidic	acid	The sentence identifies the acidic environment or conditions. Simply stating that the mechanism was an acid hydrolysis reaction does not suffice, as "acid hydrolysis" is the name of the reaction and does not itself indicate an awareness of the reaction occurring in acidic media	"Acid present in solution" "Acidic environment" "Acidic conditions"
	Specifies reaction medium—aqueous	aq	The sentence identifies the aqueous environment or conditions.	"Aqueous environment" "Water in solution" "Presence of water"
	Specifies reaction medium—body	body	The sentence identifies that the reaction is occurring in the body.	"In the body" "In the blood"
	Specifies the carbonyls involved	carb	The sentence specifies which carbonyls on the thalidomide molecule are involved in the reaction.	"Carbonyl in the 6-membered ring" "Carbonyl that is closest to the stereocenter"

				"Furthest away from the aromatic ring"
	Description of connectivity	conn	The sentence includes a depiction of the connectivity of the starting materials, intermediates, or products. This code was not applied when only the word "intermediate" was used, as simply stating that an intermediate is present gives no indication of connectivity.	"The nitrogen atom that is part of the imide group is attached to a hydrogen atom" "The Thalidomide molecule has two amide groups" "…creating a hydroxyl group" "At this moment, we have a neutral tetrahedral intermediate."
Identifying activities	Explicit electron movement	exp	The sentence uses the word "electrons" or phrase "lone pair" as the subject of a phrase when describing the movement of electrons.	"Electrons from one of the oxygens then move…" "The lone pair then comes back down to reform the double bond…"
	Implicit electron movement—entity focused	entity	The sentence uses a word or phrase other than "electrons" or "lone pair" as the subject of a phrase when describing the movement of electrons, with any verb besides "attacks."	"One of the hydroxyl substituents forms a double bond…"
	Implicit electron movement—"attacks"	att	The sentence uses a word or phrase other than "electrons" or "lone pair" as the subject of a phrase when describing the movement of electrons, with the verb "attacks."	"Water then attacks…"
	Implicit electron movement—protonates-deprotonates	prot	The sentence uses some variation of the word "protonates" or "deprotonates" to describe a mechanistic step. This code was not applied when variations of these words were used to describe a structural feature (e.g., "the protonated oxygen").	"The hydronium ion protonates…" "A water molecule deprotonates..."
	Implicit electron movement—double bond movement	dbm	The sentence refers to the movement of double bonds rather than the movement of electrons.	"This pushes the double bond up onto the oxygen…"
	Implicit electron movement—passive electron pushing	epush	The sentence uses a phrase that passively describes the movement of electrons (in the sense that the subject of the phrase is something other than the electrons or atoms/molecules involved in the mechanism).	"Electron pushing results in…" "The oxygen in the water molecule then attacks the carbon in the carbonyl, which, through electron pushing, forms a

				tetrahedral intermediate…"
Identifying properties of entities	Changes in bonding—bond breaking and making	bbm	The sentence uses language to indicate that bonds are being broken or formed in the process of a mechanistic step.	"The bond between the nitrogen and carbon breaks" "A lone pair from the oxygen reforms the carbonyl double bond."
	Changes in bonding—ring opening	ring	The sentence explicitly describes thalidomide's ring structure being broken or opened in the mechanism.	"The ring then opens" "Breaking the ring"
	Changes in bonding—nitrogen leaving	nitro	The sentence explicitly refers to the nitrogen-carbon bond breaking as the nitrogen acting as a leaving group.	"Eliminates the nitrogen" "Kicking out the nitrogen" "The nitrogen group leaves"
	Acid-base	ab	The sentence refers to a reactant acting as an acid or a base or refers to a mechanistic step as an acid-base reaction. This code was not applied when the phrase "acid hydrolysis" appeared; students needed to have included language relating to acid-base chemistry in connection to entities acting in the mechanism.	"An acid protonates…" "The carbonyl group will then be deprotonated by the conjugate base of the original acid…" "…either carbonyls are protonated through an acid/base reaction…"
	Nucleophile-electrophile	nuc	The sentence refers to the identify of reacting species as nucleophiles or electrophiles when describing a mechanistic step.	"Then, water, acting as a nucleophile, attacks the electrophilic carbon" "Electrophilic means it is extremely attracted to electrons."
	Charge	charge	The sentence refers to the creation or neutralization of formal charges when describing a mechanistic step.	"The oxygen is then deprotonated to neutralize the charge…" "The water would attack that positively charged carbonyl group." "The positive oxygen activates the carbonyl making the carbon a partial positive."
	Resonance	res	The sentence justifies a mechanistic step by referring to the resonance structures of the reacting molecules.	"The positive charge on the oxygen atom is stabilized through resonance" "The resonance form of this molecule results in a positive charge…"

				"The electrons from the double bond resonate onto the oxygen"
	Electronegativity	eneg	The sentence justifies a mechanistic step by referring to the electronegativity of the reacting atoms.	"…because nitrogen is more electronegative, the lone pair falls on the nitrogen atom" "This increases the net inductive effect on the associated carbonyl carbon since it makes the oxygen more electron deficient."

8.11.3 Appendix 3. Sample responses and application of coding scheme

Sample student response:

Thalidomide undergoes acid hydrolysis through a series of steps. In the first step, an acid protonates a water molecule to form a hydronium ion. Next, the hydronium ion protonates one of the carbonyls on the ring in thalidomide. This allows the oxygen to form an unstable positive charge. Then, water, acting as a nucleophile, attacks the electrophilic carbon. The oxygen's positive charge becomes neutral by deprotonation. The nitrogen on the 6 membered ring gets protonated by another hydronium ion and so, it becomes a good leaving group. An oxygen in one of the attached -OH groups moves one of its lone pair electrons to form a double bond. This step breaks the bond attached to the amine. Lastly, a water molecule deprotonates the positive oxygen and the molecule becomes neutral. Thalidomide can form two sterically different products because either carbonyl can be attacked.

Codes present in this response:

Describing the target phenomenon
- Overview of hydrolysis
- Identifies two reaction pathways

Identifying setup conditions
- Description of connectivity

Identifying activities
- Protonates-deprotonates
- "Attacks"
- Nitrogen leaving
- Explicit electron movement
- Bond breaking and making

Identifying properties of entities
- Acid-base
- Charge
- Nucleophile-electrophile

Sample student response:

Acid Hydrolysis can occur in the body with the following mechanism. In acidic conditions, thalidomide is protonated at the oxygen of either carbonyl, creating a formal positive charge on the oxygen. Water (H_2O) then attacks that carbonyl center and the electrons from the carbonyl move toward the oxygen to get rid of the positive charge there. After deprotonation, there are two –OH groups at the carbon that was originally attacked. The Nitrogen is then protonated and the electrons from one of these –OH groups collapses to reform original carbonyl and to kick off the –NH2 group, breaking the ring. The hydrogen bonded to the oxygen that collapsed is then deprotonated to form either of the hydrolysis products depending on which carbonyl was original attacked by the water.

Codes present in this response:

Describing the target phenomenon
- Overview of hydrolysis
- Identifies two reaction pathways

Identifying setup conditions
- Reaction medium—body
- Reaction medium—acidic
- Description of connectivity

Identifying activities
- Protonates-deprotonates
- "Attacks"
- Explicit electron movement
- Bond breaking and making
- Nitrogen leaving
- Ring opening

Identifying properties of entities
- Charge

Figure 8.14 Two example student responses, with the applied codes indicated. Note that (1) these are excerpts of the full responses, including only the portion of the response that was analyzed and (2) codes were applied on the sentence level, and have been indicated on a finer grain size to demonstrate the portions of each sentence that correspond to the applied codes.

8.11.4 Appendix 4. Appearance rate and frequency data

Table 8.2 Appearance rates and frequency data for each category and code. Entries without frequency data or descriptive statistics are the categories for which only sub-codes were applied. To contextualize this data, note that the average response contained 9.81 sentences (with standard deviation 2.55 sentences) and had 22.25 codes applied (with standard deviation 6.26 codes)

Category/code	Appearance[a] (%)	Frequency[b]	Mean[c]	St. dev.[c]
Describing the target phenomenon	99			
Overview of hydrolysis	98	402	2.51	1.20
Identifies two reaction pathways	86	214	1.52	0.67
Identifying setup conditions	96			
Specifies reaction medium	74			
Acidic	50	133	1.62	0.87
Aqueous	29	59	1.23	0.51
Body	42	88	1.29	0.62
Specifies the carbonyls involved	55	132	1.47	0.69
Description of connectivity	82	274	2.04	1.21
Identifying activities	100			
Describes electron movement	99			
Explicit electron movement	85	263	1.88	0.84
Implicit electron movement	99			
Entity focused	18	37	1.23	0.50
“Attacks”	90	205	1.40	0.65
Protonates-deprotonates	96	581	3.72	1.22
Double bond movement	6	9	1.00	0.00
Passive electron pushing	1	2	1.00	0.00
Describes changes in bonding	100			
Bond breaking and making	82	202	1.52	0.78
Ring opening	48	85	1.08	0.27
Nitrogen leaving	61	132	1.33	0.55
Identifying properties of entities	95			
Acid-base	67	233	2.14	1.16
Nucleophile-electrophile	55	143	1.61	0.86
Charge	83	414	3.04	1.54
Resonance	8	15	1.15	0.38
Electronegativity	1	4	2.00	1.41

[a] Percent of responses in which the code, or any code within the category, appears at least once ($N = 163$ responses). [b] Number of sentences to which the code was applied ($N = 1497$ sentences). [c] Statistic for the frequencies, across the set of responses in which the code appeared.

8.11.5 Appendix 5. Co-occurrence and lift data

	over	idpath	acid	aq	body	carb	conn	exp	entity	att	prot	dbm	epush	bbm	ring	nitro	ab	nuc	charge	res	eneg
over	402	90	41	35	58	36	22	4	0	18	54	0	0	30	11	2	24	25	17	0	0
idpath		214	14	2	4	69	20	4	1	33	87	0	0	5	4	2	21	13	11	0	0
acid			133	16	23	17	9	11	0	15	76	0	0	9	2	8	64	13	39	0	0
aq				59	15	3	7	2	1	5	16	0	0	15	0	1	14	8	6	0	0
body					88	7	3	4	0	8	17	0	0	1	0	0	12	6	2	0	0
carb						132	8	5	3	21	79	0	0	0	2	0	28	13	13	1	0
conn							274	60	8	43	101	2	1	40	16	18	32	22	83	3	0
exp								263	4	63	37	2	1	115	35	66	21	30	96	9	2
entity									37	0	4	1	0	11	4	10	7	1	8	2	0
att										205	40	2	0	10	4	3	29	81	61	4	1
prot											581	0	0	36	14	47	194	31	239	2	0
dbm												9	0	1	0	0	0	1	4	0	0
epush													2	0	0	0	0	0	1	0	0
bbm														202	39	54	16	13	48	6	1
ring															85	32	3	2	11	2	0
nitro																132	16	4	38	3	0
ab																	233	16	101	0	0
nuc																		143	33	5	1
charge																			414	5	0
res																				15	0
eneg																					4

Figure 8.15 Co-occurrence frequency data for all codes. The values indicate the total number of sentences for which each pair of codes appeared together.

	idpath	acid	aq	body	carb	conn	exp	entity	att	prot	dbm	epush	bbm	ring	nitro	ab	nuc	charge	res	eneg
over	1.57	1.15	2.21	2.45	1.02	0.30	0.06	0.00	0.33	0.35	0.00	0.00	0.55	0.48	0.06	0.38	0.65	0.15	0.00	0.00
idpath		0.74	0.24	0.32	3.66	0.51	0.11	0.19	1.13	1.05	0.00	0.00	0.17	0.33	0.11	0.63	0.64	0.19	0.00	0.00
acid			3.05	2.94	1.45	0.37	0.47	0.00	0.82	1.47	0.00	0.00	0.50	0.26	0.68	3.09	1.02	1.06	0.00	0.00
aq				4.32	0.58	0.65	0.19	0.69	0.62	0.70	0.00	0.00	1.88	0.00	0.19	1.52	1.42	0.37	0.00	0.00
body					0.90	0.19	0.26	0.00	0.66	0.50	0.00	0.00	0.08	0.00	0.00	0.88	0.71	0.08	0.00	0.00
carb						0.33	0.22	0.92	1.16	1.54	0.00	0.00	0.00	0.27	0.00	1.36	1.03	0.36	0.76	0.00
conn							1.25	1.18	1.15	0.95	1.21	2.73	1.08	1.03	0.75	0.75	0.84	1.10	1.09	0.00
exp								0.62	1.75	0.36	1.26	2.85	3.24	2.34	2.85	0.51	1.19	1.32	3.42	2.85
entity									0.00	0.28	4.50	0.00	2.20	1.90	3.07	1.22	0.28	0.78	5.39	0.00
att										0.50	1.62	0.00	0.36	0.34	0.17	0.91	4.14	1.08	1.95	1.83
prot											0.00	0.00	0.46	0.42	0.92	2.15	0.56	1.49	0.34	0.00
dbm												0.00	0.82	0.00	0.00	0.00	1.16	1.61	0.00	0.00
epush													0.00	0.00	0.00	0.00	0.00	1.81	0.00	0.00
bbm														3.40	3.03	0.51	0.67	0.86	2.96	1.85
ring															4.27	0.23	0.25	0.47	2.35	0.00
nitro																0.78	0.32	1.04	2.27	0.00
ab																	0.72	1.57	0.00	0.00
nuc																		0.83	3.49	2.62
charge																			1.21	0.00
res																				0.00

Figure 8.16 Lift values for each pair of codes.

8.12 Acknowledgements

The authors would like to thank the Keck Foundation and the University of Michigan Third Century Initiative for funding. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1256260. The authors would additionally like to thank Solaire Finkenstaedt-Quinn, other members of the Shultz group, and Arthur Miranda for discussions related to the preparation of this manuscript.

8.13 References

(1) Anderson, T. L.; Bodner, G. M. What Can We Do about "Parker"? A Case Study of a Good Student Who Didn't "get" Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–101. https://doi.org/10.1039/b806223b.

(2) Grove, N. P.; Bretz, S. L. A Continuum of Learning: From Rote Memorization to Meaningful Learning in Organic Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 201–208. https://doi.org/10.1039/c1rp90069b.

(3) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable Reasoning: Multi-Variate Problem-Solving in Organic Chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281–292. https://doi.org/10.1039/c0rp90003f.

(4) Graulich, N. The Tip of the Iceberg in Organic Chemistry Classes: How Do Students Deal with the Invisible? *Chem. Educ. Res. Pract.* **2015**, *16* (1), 9–21. https://doi.org/10.1039/c4rp00165f.

(5) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402–1407. https://doi.org/10.1021/ed082p1402.

(6) Ferguson, R.; Bodner, G. M. Making Sense of the Arrow-Pushing Formalism among Chemistry Majors Enrolled in Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 102–113. https://doi.org/10.1039/b806225k.

(7) Grove, N. P.; Cooper, M. M.; Cox, E. L. Does Mechanistic Thinking Improve Student Success in Organic Chemistry? *J. Chem. Educ.* **2012**, *89* (7), 850–853. https://doi.org/10.1021/ed200394d.

(8) Grove, N. P.; Cooper, M. M.; Rush, K. M. Decorating with Arrows: Toward the Development of Representational Competence in Organic Chemistry. *J. Chem. Educ.* **2012**, *89* (7), 844–849. https://doi.org/10.1021/ed2003934.

(9) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Ideas about Nucleophiles and Electrophiles: The Role of Charges and Mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797–810. https://doi.org/10.1039/c5rp00113g.

(10) Cooper, M. M.; Kouyoumdjian, H.; Underwood, S. M. Investigating Students' Reasoning about Acid-Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(11) Bhattacharyya, G.; Harris, M. S. Compromised Structures: Verbal Descriptions of Mechanism Diagrams. *J. Chem. Educ.* **2018**, *95* (3), 366–375. https://doi.org/10.1021/acs.jchemed.7b00157.

(12) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.;

Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions via App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 878–892. https://doi.org/10.1039/c9rp00260j.

(13) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chem. Educ. Res. Pract.* **2018**, *19* (4), 1117–1141. https://doi.org/10.1039/c8rp00131f.

(14) Crandell, O. M.; Kouyoumdjian, H.; Underwood, S. M.; Cooper, M. M. Reasoning about Reactions in Organic Chemistry: Starting It in General Chemistry. *J. Chem. Educ.* **2019**, *96* (2), 213–226. https://doi.org/10.1021/acs.jchemed.8b00784.

(15) Grove, N. P.; Hershberger, J. W.; Bretz, S. L. Impact of a Spiral Organic Curriculum on Student Attrition and Learning. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 157–162. https://doi.org/10.1039/b806232n.

(16) 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.

(17) 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.

(18) Galloway, K. R.; Stoyanovich, C.; Flynn, A. B. Students' Interpretations of Mechanistic Language in Organic Chemistry before Learning Reactions. *Chem. Educ. Res. Pract.* **2017**, *18* (2), 353–374. https://doi.org/10.1039/c6rp00231e.

(19) Webber, D. M.; Flynn, A. B. How Are Students Solving Familiar and Unfamiliar Organic Chemistry Mechanism Questions in a New Curriculum? *J. Chem. Educ.* **2018**, *95* (9), 1451–1467. https://doi.org/10.1021/acs.jchemed.8b00158.

(20) Grimberg, B. I.; Hand, B. Cognitive Pathways: Analysis of Students' Written Texts for Science Understanding. *Int. J. Sci. Educ.* **2009**, *31* (4), 503–521. https://doi.org/10.1080/09500690701704805.

(21) Moreira, P.; Marzabal, A.; Talanquer, V. Using a Mechanistic Framework to Characterise Chemistry Students' Reasoning in Written Explanations. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 120–131. https://doi.org/10.1039/c8rp00159f.

(22) Moon, A.; Moeller, R.; Gere, A. R.; Shultz, G. V. Application and Testing of a Framework for Characterizing the Quality of Scientific Reasoning in Chemistry Students' Writing on Ocean Acidification. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 484–494. https://doi.org/10.1039/c9rp00005d.

(23) Reynolds, J. A.; Thaiss, C.; Katkin, W.; Thompson, R. J. Writing-to-Learn in Undergraduate Science Education: A Community-Based, Conceptually Driven Approach. *CBE Life Sci. Educ.* **2012**, *11* (1), 17–25. https://doi.org/10.1187/cbe.11-08-0064.

(24) Shultz, G. V.; Gere, A. R. Writing-to-Learn the Nature of Science in the Context of the Lewis Dot Structure Model. *J. Chem. Educ.* **2015**, *92* (8), 1325–1329. https://doi.org/10.1021/acs.jchemed.5b00064.

(25) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Chambers, T. G.; Moon, A.; Goldman, R. S.; Gere, A. R.; Shultz, G. V. Investigation of the Influence of a Writing-To-Learn Assignment on Student Understanding of Polymer Properties. *J. Chem. Educ.* **2017**, *94* (11), 1610–1617. https://doi.org/10.1021/acs.jchemed.7b00363.

(26) Moon, A.; Zotos, E.; Finkenstaedt-Quinn, S.; Gere, A. R.; Shultz, G. Investigation of the Role of Writing-to-Learn in Promoting Student Understanding of Light-Matter Interactions. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 807–818. https://doi.org/10.1039/c8rp00090e.

(27) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and Conceptual Learning in Science: An Analysis of Assignments. *Writ. Commun.* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(28) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(29) Bhattacharyya, G. From Source to Sink: Mechanistic Reasoning Using the Electron-Pushing Formalism. *J. Chem. Educ.* **2013**, *90* (10), 1282–1289. https://doi.org/10.1021/ed300765k.

(30) Becker, N.; Noyes, K.; Cooper, M. Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *J. Chem. Educ.* **2016**, *93* (10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298.

(31) Weinrich, M. L.; Talanquer, V. Mapping Students' Modes of Reasoning When Thinking about Chemical Reactions Used to Make a Desired Product. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 394–406. https://doi.org/10.1039/c5rp00208g.

(32) Caspari, I.; Weinrich, M. L.; Sevian, H.; Graulich, N. This Mechanistic Step Is "Productive": Organic Chemistry Students' Backward-Oriented Reasoning. *Chem. Educ. Res. Pract.* **2018**, *19* (1), 42–59. https://doi.org/10.1039/c7rp00124j.

(33) Cartrette, D. P.; Mayo, P. M. Students' Understanding of Acids/Bases in Organic Chemistry Contexts. *Chem. Educ. Res. Pract.* **2011**, *12* (1), 29–39. https://doi.org/10.1039/c1rp90005f.

(34) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Fragmented Ideas about the

Structure and Function of Nucleophiles and Electrophiles: A Concept Map Analysis. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 1019–1029. https://doi.org/10.1039/c6rp00111d.

(35) Finkenstaedt-Quinn, S. A.; Watts, F. M.; Petterson, M. N.; Archer, S. R.; Snyder-White, E. P.; Shultz, G. V. Exploring Student Thinking about Addition Reactions. *J. Chem. Educ.* **2020**, *97* (7), 1852–1862. https://doi.org/10.1021/acs.jchemed.0c00141.

(36) Christian, K.; Talanquer, V. Modes of Reasoning in Self-Initiated Study Groups in Chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 286–295. https://doi.org/10.1039/c2rp20010d.

(37) 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.

(38) 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.

(39) Cruz-Ramírez De Arellano, D.; Towns, M. H. Students' Understanding of Alkyl Halide Reactions in Undergraduate Organic Chemistry. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 501–515. https://doi.org/10.1039/c3rp00089c.

(40) Popova, M.; Bretz, S. L. "It's Only the Major Product That We Care about in Organic Chemistry": An Analysis of Students' Annotations of Reaction Coordinate Diagrams. *J. Chem. Educ.* **2018**, *95* (7), 1086–1093. https://doi.org/10.1021/acs.jchemed.8b00153.

(41) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Res. Teach. English* **2015**, *50* (2), 199–235.

(42) Finkenstaedt-Quinn, S. A.; Snyder-White, E. P.; Connor, M. C.; Gere, A. R.; Shultz, G. V. Characterizing Peer Review Comments and Revision from a Writing-to-Learn Assignment Focused on Lewis Structures. *J. Chem. Educ.* **2019**, *96* (2), 227–237. https://doi.org/10.1021/acs.jchemed.8b00711.

(43) Finkenstaedt-Quinn, S. A.; Halim, A. S.; Kasner, G.; Wilhelm, C. A.; Moon, A.; Gere, A. R.; Shultz, G. V. Capturing Student Conceptions of Thermodynamics and Kinetics Using Writing. *Chem. Educ. Res. Pract.* **2020**, *21*, 922–939. https://doi.org/10.1039/C9RP00292H.

(44) Emig, J. Writing as a Mode of Learning. *Coll. Compos. Commun.* **1977**, *28* (2), 122–128.

(45) Klein, P. D. Reopening Inquiry into Cognitive Processes in Writing-To-Learn. *Educ. Psychol. Rev.* **1999**, *11* (3), 203–270. https://doi.org/10.1023/A:1021913217147.

(46) MacArthur, C. A.; Graham, S. Writing Research from a Cognitive Perspective. In *Handbook of Writing Research*; MacArthur, C. A., Graham, S., Fitzgerald, J., Eds.; Guilford: New York, 2016; Vol. 40, pp 24–40.

(47) Klein, P. D.; Boscolo, P. Trends in Research on Writing as a Learning Activity. *J. Writ. Res.* **2016**, *7* (3), 311–351. https://doi.org/10.17239/jowr-2016.07.03.01.

(48) Flower, L.; Hayes, J. R. A Cognitive Process Theory of Writing. *Coll. Compos. Commun.* **1981**, *32* (4), 365–387.

(49) Flower, L.; Hayes, J. R. Images, Plans, and Prose: The Representation of Meaning in Writing. *Writ. Commun.* **1984**, *1* (1), 120–160.

(50) Hayes, J. R. A New Framework for Understanding Cognition and Affect in Writing. In *The Science of Writing: Theories, Methods, Individual Differences, and Applications*; Levy, C. M., Ransdell, S., Eds.; Lawrence Erbaum Associates: Mahwah, New Jersey, 1996; pp 1–27.

(51) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Sci. Educ.* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(52) Machamer, P.; Darden, L.; Craver, C. F. Thinking about Mechanisms. *Philos. Sci.* **2000**, *67* (1), 1–25.

(53) Corbin, J. M.; Strauss, A. Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qual. Sociol.* **1990**, *13* (1), 3–21. https://doi.org/10.1007/BF00988593.

(54) Nowell, L. S.; Norris, J. M.; White, D. E.; Moules, N. J. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *Int. J. Qual. Methods* **2017**, *16* (1), 1–13. https://doi.org/10.1177/1609406917733847.

(55) Miles, M. B.; Huberman, A. M.; Saldana, J. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed.; Sage: Los Angeles, CA, 2014.

(56) Kirilenko, A. P.; Stepchenkova, S. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLoS One* **2016**, *11* (3), 1–14. https://doi.org/10.1371/journal.pone.0149787.

(57) McHugh, M. L. Lessons in Biostatistics Interrater Reliability : The Kappa Statistic. *Biochem. Medica* **2012**, *22* (3), 276–282.

(58) QSR International Pty Ltd. NVivo Qualitative Data Analysis Software (Version 12). 2018.

(59) RStudio Team. RStudio: Integrated Development for R. 2018.

(60) Merceron, A.; Yacef, K. Interestingness Measures for Association Rules in Educational

Data. *Educ. Data Min. 2008 - 1st Int. Conf. Educ. Data Mining, Proc.* **2008**, 57–66.

(61) Graulich, N.; Bhattacharyya, G. Investigating Students' Similarity Judgments in Organic Chemistry. *Chem. Educ. Res. Pract.* **2017**, *18* (4), 774–784. https://doi.org/10.1039/c7rp00055c.

(62) Dood, A. J.; Fields, K. B.; Raker, J. R. Using Lexical Analysis to Predict Lewis Acid-Base Model Use in Responses to an Acid-Base Proton-Transfer Reaction. *J. Chem. Educ.* **2018**, *95* (8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177.

Chapter 9

Developing Machine Learning Models for Automatic Analysis of Organic Chemistry Students' Written Descriptions of Organic Reaction Mechanisms

9.1 Initial remarks

This chapter focuses on analyzing students' responses to organic chemistry writing-to-learn (WTL) assignments using machine learning (ML) methods for automated text analysis, presenting a transition into the final portion of the dissertation. The chapter specifically builds on the framework and analysis presented in Chapter 8 by extending the analysis to three different reaction mechanisms in organic chemistry. Extending the framework allows for identifying students' mechanistic reasoning within students' descriptions and explanations of reaction mechanisms more broadly. Furthermore, the chapter presents the development of ML models that can automatically analyze students' responses to the WTL assignments and identify the presence of various features of students' mechanistic reasoning at the sentence level. These models are the basis for the automated formative assessment tool described in Chapter 10. This research contributes to the literature within STEM education research on using ML methodologies to characterize students' written responses to formative assessment items by specifically demonstrating how these methods can be applied to the longer, essay-length responses elicited by WTL assignments.

The chapter describes the analysis of students' responses to two of the WTL assignments administered in the second-semester organic chemistry laboratory course. Students' responses to the two assignments include their descriptions and explanations of three different reaction mechanisms: an acid-catalyzed amide hydrolysis reaction (the same reaction discussed in Chapter 8), a racemization reaction, and a base-free Wittig reaction. Students' responses were analyzed using the framework described in Chapter 8, which was further developed in this work to allow for broader application to the three different mechanisms. Results from the study highlight the features of students' mechanistic reasoning in their writing about the three mechanisms, demonstrating how the features students include differ depending on the WTL assignment and the mechanism discussed. For example, students discussed the reaction medium much more in their responses to one of the WTL assignments in comparison to their responses to the other WTL assignment, plausibly due to the prompting within the assignment description. Additionally, students discussed stereochemistry and the formation of stereochemistry much more for the racemization mechanism in comparison to the other two mechanisms. These findings demonstrate that both the prompting within the WTL assignment and the specific mechanism students are asked to describe can influence the features of students' reasoning that are elicited. The study also describes a set of predictive models developed using machine learning techniques. These models can accurately identify the features of students' writing in alignment with the analysis framework. The development of these machine learning models contributes to the new and growing area of research focused on using machine learning for formative assessments in undergraduate STEM courses.

This chapter will be published by the Royal Society of Chemistry in the forthcoming book *Student Reasoning in Organic Chemistry*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author, I contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing). A.J. Dood contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing). G.V. Shultz contributed to project supervision, conceptualization, and writing (review and editing).

9.2 Abstract

Many assessments in organic chemistry ask students to produce reaction mechanisms with the electron-pushing formalism. It is well known that students can apply the electron-pushing formalism without engaging in chemical reasoning about the processes underlying mechanisms.

Furthermore, engagement in mechanistic and causal reasoning correlates with student performance on organic chemistry tasks. Hence, it is valuable to elicit students' explanations of mechanisms beyond relying on traditional mechanism assessments. One evidence-based approach for encouraging and eliciting students' mechanistic explanations is through writing; however, instructors may hesitate to implement writing in their courses due to a lack of tools available to provide formative feedback on students' mechanistic explanations. To address this challenge, we analyzed students' written explanations of three different organic reaction mechanisms for individual features involved in mechanistic reasoning. In this chapter, we present our adaptation of Russ et al.'s mechanistic reasoning framework specifically for students' written explanations of organic chemistry reaction mechanisms. Additionally, we describe a set of predictive models which we have used to accurately identify features of students' writing involved in mechanistic reasoning in the context of the three different reaction mechanisms. This work has implications for instructors seeking to identify students' reasoning in written explanations of organic reaction mechanisms. Additionally, this work has implications for future research into developing immediate and automated student- and instructor-facing formative feedback to encourage students' development of mechanistic and causal reasoning.

9.3 Introduction

The chemistry education research literature suggests that students need instructional support to engage with the underlying chemical properties and processes when considering reaction mechanisms in organic chemistry. Recent studies have emphasized the value of having students write descriptions and explanations of reaction mechanisms to provide students with opportunities to engage in mechanistic reasoning. Cooper et al.,[1] Crandell et al.,[2,3] and Dood et al.[4–7] have all explored using constructed response or open-ended assessment items to engage students in mechanistic reasoning. A recent study by Watts et al.[8] examined writing-to-learn (WTL) as another approach for eliciting evidence of students' mechanistic reasoning through writing. These approaches lead to the inherent challenge that many organic chemistry instructors may lack the time or resources to provide constructive, formative feedback about students' writing. As such, recent attention in the chemistry education research literature has focused on developing tools to provide formative feedback on students' writing using machine learning techniques.[4–7,9–12] The goal of this work is to describe an approach using automated text analysis to identify evidence of students' mechanistic reasoning across their descriptions and explanations of different reaction mechanisms elicited through WTL assignments.

9.3.1 Eliciting students' mechanistic reasoning in organic chemistry through writing

Reaction mechanisms are central to organic chemistry, as practicing organic chemists use mechanisms to explain and predict the outcomes of reactions.[13–15] Because of the centrality of mechanisms to the discipline, eliciting and supporting students' mechanistic reasoning in organic chemistry is a common focus in organic chemistry education research. Many studies concerning mechanistic reasoning focus on students' use of the electron-pushing formalism (EPF), the curved arrow notation representing the flow of electrons in a reaction mechanism.[16–20] These studies identified that many students, particularly those enrolled in traditional organic chemistry curricula, tend to memorize components of reaction mechanisms and to use the EPF as an academic exercise not connected to physical meaning.[17–19,21,22] Other studies of students' mechanistic reasoning identify the challenges students face with connecting surface features of reactions to their implicit meaning[23–27] and with considering the various chemical and physical properties guiding reaction mechanisms.[28–33] Though challenging for students, identifying multiple implicit properties of entities and deciding which properties are relevant to the problem are important factors in students' problem-solving ability and their success in problem solving.[20,25,33–38] These studies point to the need to better support students' reasoning with reaction mechanisms—that is, the need to support students with making connections between representations of reactions, the EPF, and underlying chemical and physical properties. In efforts to achieve this goal, recent studies have elicited students' mechanistic reasoning through writing, including constructed response prompts[1–7,10,35] and WTL assignments.[8]

Cooper et al.[1] described a constructed response item that elicited students' mechanistic reasoning for an acid–base reaction. Their study found that students engaged in various types of reasoning in response to the prompt, though the nature of the prompt influenced students' responses. Specifically, different students engaged in descriptive reasoning (*what* happens), causal reasoning (*what* happens and *why*), and mechanistic reasoning (*what* happens and *how*). Crandell et al.[2] implemented a constructed response item focused on a Lewis acid–base complexation reaction, similarly finding that the prompt elicited responses including mechanistic reasoning, causal reasoning, and causal mechanistic reasoning. Crandell et al. defined causal mechanistic reasoning as encompassing *what* happens in a mechanism along with *why* (e.g., due to electrostatic interactions between reactants) and *how* (accounting for electron movement). In a follow-up study, Crandell et al.[3] administered a constructed response prompt focused on substitution reactions and similarly found that aspects of the prompt influenced students' responses and that the prompt elicited a range of students' reasoning. The studies by Cooper et al.[1] and Crandell et al.[2,3] were in the context of transformed general[39] or organic[40] chemistry courses; Dood et al.[4–7] administered constructed response items to students in a traditional organic chemistry curriculum. Dood et al.[5,6] used lexical analysis of students' short, written responses to identify trends across students' responses at scale, and in additional studies[4,7] demonstrated how automated analysis could be used to provide students with adaptive tutorials to support their learning. Yik et al.[10] used lexical analysis to identify the correct use of the Lewis acid–base model in written responses about several different reaction mechanisms. In a study using WTL as opposed to constructed response items, Watts et al.[8] examined students' responses to a writing assignment eliciting students' mechanistic reasoning about an acid hydrolysis mechanism. Their findings illustrated how the analysis of students' writing for the necessary components of mechanistic reasoning could demonstrate trends in mechanistic reasoning across students' responses. For example, the writing analysis by Watts et al.[8] identified evidence of the hierarchical nature of mechanistic reasoning and indicated the different ways students wrote about alternative pathways for the hydrolysis reaction.

9.3.2 Machine learning for analyzing student writing in chemistry

Though it is well known that students can benefit from explaining the phenomena they are learning about, and that writing assignments have the potential to make students' reasoning tractable,[1–7,9,41,42] there are also barriers to implementation of writing in courses.[43] Previous work has found that instructors believe writing is useful but feel it is too difficult to implement in their courses due to time constraints and lack of ability to provide students with feedback on their writing.[43] In an effort to ease the implementation of writing in chemistry courses, several recent studies have developed automated text analysis (ATA) models which can be used to provide students and instructors with immediate feedback on students' writing.[4–7,9–12] Specifically, ATA models can be developed for assignments that ask students to describe and explain reaction mechanisms (i.e., elicit students' mechanistic reasoning).

Previous work on ATA models in chemistry has focused on single writing prompts, [4–7,9,11] peer review comments,[12] or on a single aspect of multiple constructed-response items.[10] Though these models are indeed useful, their scope is limited. The development of a single model is time consuming (i.e., thousands of students' responses must be coded by researchers prior to an iterative process of model development), so models which can be implemented more broadly are desirable. To overcome these limitations, the goal of this study is to describe ATA models which can identify several aspects of mechanistic reasoning for students' explanations of three different organic chemistry reaction mechanisms. The models presented here also focus on the sentence level, meaning that the model evaluates each sentence separately. This approach may allow instructors and students to receive more detailed feedback than models which provide automated feedback for an entire response. To this end, the models can be useful to provide feedback on long prompts (e.g., WTL assignments), as the feedback provided is tied to a single sentence.

9.4 Theoretical framework

The theoretical framework guiding this study is the cognitive process theory of writing originally developed by Flower and Hayes[44,45] and later revised by Hayes.[46] This cognitive perspective on writing informed both the design and implementation of the WTL assignments central to this study, and additionally provided a theoretical grounding for using students' writing as evidence of their reasoning. Hayes' revised cognitive process theory includes two components: the writer and the task environment. The writer interacts with the task environment, which includes the writing assignment, the genre and audience, and the text-in-production. During writing, the writer engages in the cognitive processes of producing text, interpreting text, and reflecting. Each of these cognitive processes can occur during any stage of the traditional writing process of planning, writing, and revising; Hayes emphasizes that the cognitive processes are recursive and do not occur linearly. As the writer engages in the cognitive processes, they use their working and long-term memory to access their knowledge and represent it in writing. This writing theory supports the design and implementation of the WTL assignments, which focus on course content rather than learning to write and incorporate structures for peer review and revision to emphasize and structure the recursive nature of writing. Furthermore, the theory provides a basis for analyzing students' writing as a means of accessing their reasoning because throughout the recursive processes of writing, students must identify how to represent their knowledge in response to the

WTL assignment. Thus, the features students incorporate in their writing reflect their knowledge and what they find important for representing their reasoning.

9.5 Research questions

Following from the cognitive process theory of writing, this research recognizes writing as a representation of knowledge. Therefore, the first goal of this study is to identify common features of students' writing about different organic reaction mechanisms. The second goal of this study is to identify whether ATA models can feasibly identify these aspects of students' writing. The following research questions guided this study:

1. How do students respond to WTL assignments intended to elicit *how* and *why* organic reaction mechanisms occur?
2. Does automated text analysis allow for predictions of the components included in students' written mechanistic descriptions?

9.6 Methods

9.6.1 Setting and participants

This research was conducted at a large, research-intensive university in the Midwestern United States. The course setting was a second-semester introductory organic chemistry laboratory course, and the participants were the 771 students who received a final grade in the course. The students may have had previous experience with WTL assignments, though data was not collected concerning prior experiences with WTL. Students in the course are primarily non-chemistry majors, with majors including neuroscience and biomolecular science. The laboratory course is offered separately from the second-semester organic chemistry lecture course, though students often take the courses concurrently. The laboratory course includes a 1 h lecture taught by either faculty or lecturers and a 4 h laboratory taught by graduate student instructors. The lecture portion of the course covers topics related to the laboratory sessions, including specific reactions, spectral interpretation, and thinking chemically about what is happening in the laboratory. Students usually take the laboratory course in their first or second year at the university, followed by courses in analytical, inorganic, and physical chemistry in later years for chemistry majors.

9.6.2 Writing-to-learn assignments and implementation

The WTL assignments in the course were designed and implemented to support students' conceptual learning, following the structure outlined by Finkenstaedt-Quinn et al.[47] Two of the three WTL assignments administered in the laboratory course were used for this study because of their focus on eliciting descriptions and explanations of reaction mechanisms. Both assignments incorporated a meaning-making task in alignment with the literature on components of successful WTL assignments.[48,49] The meaning-making tasks required students to apply their content knowledge of organic chemistry and the principles of reaction mechanisms to the tasks of describing and explaining unfamiliar mechanisms of reactions.[50] One of the assignments elicited students' descriptions and explanations of two mechanisms that affect the thalidomide molecule: racemization and acid hydrolysis. The other assignment elicited students' descriptions and explanations of a base-free Wittig reaction. The assignments will hereafter be referred to as the thalidomide and Wittig assignments, respectively.

The mechanistic descriptions and explanations analyzed were a subset of the full-length assignments, typically ranging in length from one to two paragraphs. The relevant parts of the prompts are presented here, while the full text of the assignments can be found in the 9.11 Appendix. The relevant portion of the thalidomide prompt stated:

> "Provide thorough descriptions of the mechanisms of both racemization and acid hydrolysis, highlighting the critical structural features of thalidomide and their role in these mechanisms.
>
> a. When racemization occurs, what changes occur in the molecule?
> b. When hydrolysis occurs, what changes occur in the molecule?"

The relevant portion of the Wittig prompt stated:

> "Explain the critical structural and electronic features and properties of the starting materials and reagents in [the base-free Wittig reaction] and their role in the mechanistic steps that lead to the formation of the products without the use of an external base.
>
> a. In describing the mechanistic steps for the reaction in [the base-free Wittig reaction], what changes occur within those steps to the starting materials and reagents that lead to the formation of the ylide? (Note that the ylide is not shown in this scheme.)
> b. What structural changes happen to PBu_3 at each mechanistic step? Focus on the *how* and *why* as well as the *what.*"

The WTL assignments were implemented in the course with structures intended to support students' success: peer review, revision, and writing fellows. After having one week to submit their initial drafts, students underwent a required peer review process in which they had three days to provide content-focused feedback to typically three peers. The peer review process was double-blind, and the assignment of peer reviewers was automated using a tool in the course's online learning system. The content-focused peer review rubrics are included in the 9.11 Appendix. After peer review, students were given four days to revise their assignments based on the feedback from their peers. While the assignment was ongoing, students had access to writing fellows, undergraduates previously successful in the course who participated in a seminar course focused on preparing them to support students on the WTL assignments. The process of peer review and revision, alongside the availability of writing fellows, provided sources for interactive writing processes, metacognition, and reflection, all of which are components of successful WTL assignments.[48,49] Further details behind the general development and implementation of the WTL assignments can be found in Finkenstaedt-Quinn et al.[47]

9.6.3 Data collection

The data collected for this research includes students' final drafts in response to the two WTL assignments described above. In alignment with the cognitive process theory of writing, only final drafts were analyzed so as to best capture students' understanding after engaging in the writing process supported by the WTL implementation.[46] All data collection was approved by the Institutional Review Board for human subjects research, and all students whose data were collected consented to participate in the study. In total, 40 responses to the thalidomide WTL assignment and 40 responses to the Wittig WTL assignment were randomly selected for inclusion in the analysis (120 individual mechanistic descriptions and a total of 1243 sentences). We determined that this number of assignments represented the codes used well enough to use as a starting point for model development with the potential to add additional data to better represent specific codes.

9.6.4 Data analysis

Analytical framework. Students' final drafts for the two WTL assignments were analyzed in alignment with the framework introduced by Russ et al.[51] for identifying mechanistic reasoning in students' discourse. The framework is based on the philosophy of science literature that provides generalized descriptions of mechanisms.[52,53] These descriptions are centered around entities (the individual components of mechanisms), the properties of entities, and activities (the interactions between entities that produce change). For organic chemistry reaction mechanisms, entities are the electrons, atoms, and molecules, while activities are the electron movements resulting in bonding changes. Previous research of students' mechanistic reasoning in organic chemistry has utilized this framework.[8,37,38,54] The present study builds upon the coding scheme described by Watts et al.,[8] which was used to characterize the features necessary for mechanistic reasoning present in students' descriptions of an acid hydrolysis mechanism (which were elicited from an earlier implementation of the thalidomide WTL prompt used in this study). The goal of expanding the coding scheme for this study is to encompass multiple reaction types and improve the coding scheme's general applicability. Beyond modifying the scheme used by Watts et al., the coding scheme was also adapted to include a category to capture students' causal reasoning, a common feature of students' mechanistic reasoning described in the organic chemistry education research literature and observed in our data.[1,3,29] The central goal of the analysis was to identify features necessary for mechanistic reasoning across students' responses for the three different reaction mechanisms. Guided by the cognitive process theory of writing, we posit that the features students include after the process of writing and revising indicate what students find important for representing their reasoning. The full coding scheme and details of the analysis can be found in the results and discussion section.

Reliability. The analysis and development of the coding scheme involved meeting frequently and discussing the definitions, examples, and applications of the coding scheme. Developing and applying the coding scheme took place across multiple stages in which the first two authors independently coded students' writing, met to discuss the application of codes, and determined measures of inter-rater reliability (IRR). The codes were applied to students' responses at a sentence level, and multiple codes could be applied to each sentence. Since multiple codes could be applied to each sentence, percent agreement and fuzzy kappa were calculated at each stage.[55] Across the coding stages, the percent agreement ranged from 69% to 79% and the fuzzy kappa ranged from 0.62 to 0.74. As these values reflect moderate agreement, we discussed all disagreements and achieved consensus for the final application of codes for each unit of analysis.[56] Since the coding scheme was modified and adjusted in the early stages of analysis, we returned to the earlier-coded data to ensure consistent application of codes across the dataset.

Development of automated text analysis models. The ATA model development and evaluation was performed using the Scikit-learn and Keras libraries in Python 3.[57–59] To train the machine learning models, the data were split into a training set (used to train the model), a validation set (used to test and refine the model during training), and a testing set (i.e., set of data kept out of the training set for testing the model on unseen data). Thus, 67.5% of the data were used for training, 22.5% for validation, and 10% for testing. Several common machine learning models were tested for each aspect, including traditional models (i.e., naive Bayes, logistic regression, support vector machines) and deeper learning models (i.e., convolutional neural networks, transformer models). The models were evaluated using accuracy, Cohen's kappa,[60] Matthews correlation coefficient (MCC),[61] and true and false positives and negatives. Convolutional neural networks (CNNs) were the best overall models for each feature, taking into account the variety of fit statistics.[62,63] Once it was determined that CNNs were the appropriate model type, each model was tested with different optimizers, word embeddings, and preprocessing techniques. After this optimization, some of the models were still performing unsatisfactorily, likely due to an imbalance in the coded data (i.e., an uneven number of positive *versus* negative instances). To remedy this, additional sentences drawn from students' writing were coded by the researchers to be added to the models for specific features of mechanistic reasoning. As there were fewer positive instances for all mechanistic reasoning features, only additional positive instances were added to the data. Once optimized, each model was run 30 times and performance metrics were calculated based on an average of all 30 runs.

9.7 Results and discussion

9.7.1 How do students respond to WTL assignments intended to elicit how and why organic reaction mechanisms occur?

Building on the mechanistic reasoning framework developed by Russ et al.[51] and adapted by Watts et al.,[8] our analysis identified several features of mechanistic and causal reasoning across students' responses for the three different reaction mechanisms. There are six categories in the final coding scheme: identifying setup conditions, identifying explicit properties, identifying activities, identifying implicit properties, organization of entities, and causal reasoning. These categories are derived from the Russ et al. framework and contain eleven different codes that are specific to organic reaction mechanisms (full coding scheme in Table 9.1), focusing on implicit *versus* explicit as well as static *versus* dynamic descriptions (Figure 9.1). The codes within each category are derived from the Watts et al. framework. Many of the static and dynamic codes correspond to static *versus* dynamic descriptions of the same phenomena. For example, the static code "stereochemistry" indicates students' identification of stereochemistry in a molecule, while the dynamic code "formation of stereochemistry" indicates students' dynamic descriptions of the formation of stereochemistry. Further examples of the codes applied to sentences from students' writing are provided in Figure 9.1.

Table 9.1 The coding scheme, including definitions and examples, for identifying features of students' mechanistic reasoning. The categories provide conceptual organization for the codes in alignment with the analytical framework. While codes were applied at the sentence level, italics have been added to the examples to indicate the portion of the text corresponding to the code. The last three columns show the percent of students who included each feature in their writing per mechanism. For each mechanism, there are 40 total students whose writing was analyzed. Hyd. = thalidomide hydrolysis mechanism, Rac. = thalidomide racemization mechanism, Wittig = Wittig mechanism

Category	**Code**	**Definition**	**Example**	**Hyd. (%)**	**Rac. (%)**	**Wittig (%)**
Identifying setup conditions	Reaction medium	The sentence identifies aspects of the reaction medium (e.g., acidic or basic conditions)	"*In the acidic conditions of the stomach*, either carbonyl group…"	70.0	65.0	2.5
Identifying explicit properties of entities	Connectivity of molecules	The sentence identifies a specific description of what atom is being referred to by referring to the connections to other atoms in the molecule	"Thalidomide also contains two secondary *amides connected by the same nitrogen group* (NH)."	87.5	92.5	95.0
	Charges	The sentence identifies the formal charge (or neutrality) of atoms or molecules	"..the *negatively charged carbon* adds to the carbonyl..."	92.5	67.5	95.0
	Stereochemistry	The sentence identifies the stereochemical, regiochemical, or spatial orientation of a molecule	"The two different 3D orientations of the atoms surrounding the chiral center are called *R and S* stereocenters…"	25.0	100.0	25.0
Identifying activities	Electron movement	The sentence identifies the movement of electrons or "lone pairs"	"The *electrons from the oxygen* in the alcohol group *shift down*	82.5	80.0	92.5

			to form…”			
	Non-electronic mechanism	The sentence identifies a mechanistic activity without describing electron movement or specifying bonds being broken/made	“The carbon now *attacks* the carbon in the carbonyl group…”	100.0	97.5	100.0
	Bond breaking and making	The sentence identifies the changes in bonding during a mechanistic step	“...a whole *water molecule bonds to the carbon* attached to the OH…”	97.5	80.0	97.5
Identifying implicit properties of entities	Implicit properties	The sentence identifies a specific entity and refers to it using an implicit property, including nucleophilicity/electrophilicity, acidity/basicity, partial charges, etc.	“The *acidic proton* on the carbon that is connected to the phosphorus cation gets easily deprotonated…”	72.5	60.0	85.0
Organization of entities	Formation of stereochemistry	The sentence identifies the dynamic formation of stereochemistry, regiochemistry, or spatial orientation by describing the interaction of entities relative to one another in space	“...there is no preference for the *hydrogen coming from above or below the molecule*.”	15.0	97.5	20.0
Causal reasoning	Cause–effect only	The sentence identifies a cause and effect in their description of mechanistic activities, without using electronic properties for the reasoning	“This newly formed hydronium molecule can then protonate the nucleophilic nitrogen, *creating* a positive charge on the nitrogen atom.”	90.0	97.5	97.5
	Electronic causal	The sentence identifies a cause and effect in their description of mechanistic activities, using electronic properties for the reasoning	“...the carbonyl group is imperative *because its partial negative charge* due to resonance *allows it to initially deprotonate* the hydronium ion.”	25.0	22.5	60.0

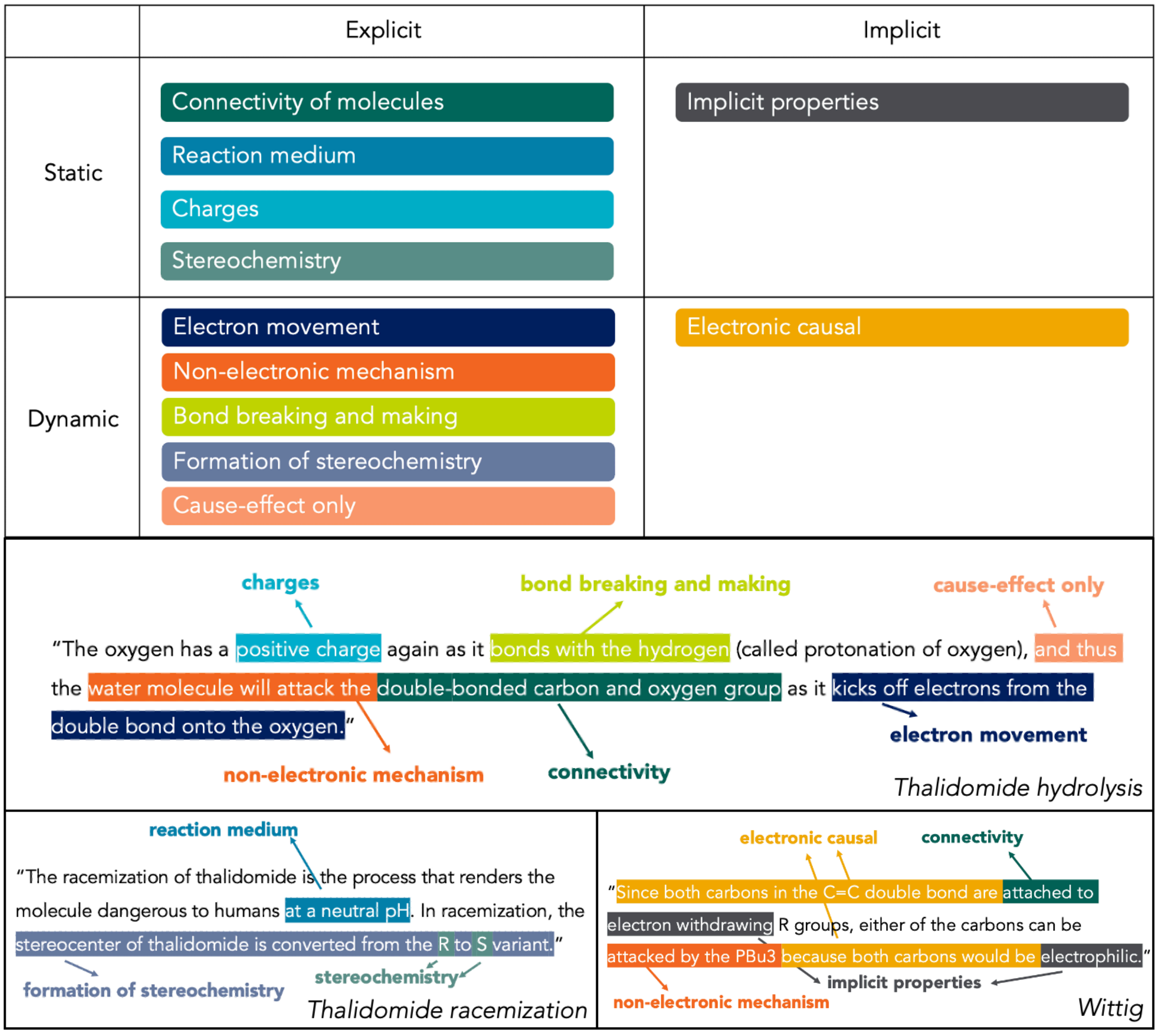


Figure 9.1 The 11 different codes separated into explicit *vs.* implicit and static *vs.* dynamic. Examples of how the codes were applied to one sentence from each of the three mechanism explanations are also provided. Note: codes were applied at the sentence level, but here we have broken them up to highlight specific parts of each sentence which refer to each code.

The identifying setup conditions category captured students' recognition of the reaction medium. Many students wrote about the reaction medium when describing the thalidomide hydrolysis and racemization mechanisms, but only one student mentioned the reaction medium when explaining the Wittig mechanism. This result is likely becuase the reaction medium (i.e., acidic conditions) is an important factor for the thalidomide reactions. The identifying explicit properties category identified students' descriptions of the static connectivity of molecules, formal charges, and stereochemistry. Across all three mechanisms, almost all students described the connectivity of molecules. This code includes explicit descriptions of how atoms are connected to other atoms in the molecule, an important part of written mechanism descriptions. Almost all students describing the thalidomide hydrolysis and Wittig mechanisms mentioned charges, while two-thirds of students mentioned charges in their description of thalidomide racemization. All students mentioned stereochemistry in their description of the thalidomide racemization mechanism, while only one-fourth of students mentioned stereochemistry when describing the other two mechanisms. Unsurprisingly, all students mentioned stereochemistry when describing a racemization reaction; a description of stereochemistry was less vital to the description and explanation of the other two mechanisms.

The identifying activities category included students' descriptions of mechanisms either electronically (i.e., explicitly referring to "electrons" or "lone pairs"), non-electronically (e.g., using terms such as "attacks" or "protonated"), or in terms of breaking and forming bonds. Most students across all three mechanisms included all three of these features. This result is promising, because identifying activities is the *what* of the reaction mechanism. Fewer students identified implicit properties; this category included one code that encompassed students' reference to implicit properties including nucleophilicity/electrophilicity, acidity/basicity, and partial charges, among others. Aligning with findings from other studies, students were more focused on explicit features, such as charges, than implicit features.[23–27] The organization of entities category included one code capturing students' descriptions of the dynamic formation of stereochemistry. Almost all students included this description when describing the thalidomide racemization mechanism. Because most of the students included both a static stereochemical description and a dynamic stereochemical description, it is promising to see that students are connecting entities (static) to activities (dynamic) in their mechanistic description of the thalidomide racemization mechanism.

Lastly, the causal reasoning category included two codes, identifying students' cause–effect reasoning (without reference to electronic properties) and students' electronic causal reasoning (i.e., causal language that connects to electronic properties). Most students across all three mechanisms included cause–effect reasoning without referencing electronic properties in their writing. Fewer students included causal reasoning employing electronic properties. This result is unsurprising given that identifying implicit properties and connecting them to the problem context is known to be challenging for students. Only 25.0% and 22.5% percent of students employed electronic causal reasoning when explaining the thalidomide hydrolysis and racemization mechanisms, respectively. Many more students (60%) employed electronic causal reasoning when explaining the Wittig mechanism. This finding suggests that the nature of the prompt may have had an influence on eliciting students' causal mechanistic reasoning. While the thalidomide prompt asked students which changes occur in the molecule and asked students to highlight critical structural features of thalidomide, the Wittig prompt asked students to explain critical structural and electronic features and properties, as well as explicitly prompted them to focus on the *why* as well as the *what*. This difference in reasoning elicited for different prompts is in line with other studies which found that the nature of the prompt influenced students' elicited reasoning.[1–3] The differences seen across mechanisms provide further evidence that all writing prompts, whether constructed-response or WTL, should be carefully constructed to elicit the desired reasoning.

9.7.2 Does automated text analysis allow for predictions of the components included in students' written mechanistic descriptions?

Using the analysis of students' writing about *how* and *why* organic reaction mechanisms occur, individual CNN models were developed to predict the presence of each feature within students' writing. For each student's response to the WTL assignment, the models can identify which features of students' mechanistic reasoning are present within each sentence. The models were developed using the iterative process described in the methods section. The performance metrics for each model are presented in Table 9.2, which includes accuracy, Cohen's kappa, MCC, and the confusion matrix. All performance metrics are calculated by using the developed model to predict the presence of each feature on the testing set (the 10% of human-coded data that was not used in model development) and then comparing the computer predictions to the human codes. The accuracy value reflects the proportion of correct computer predictions across the testing set. Cohen's kappa is a metric that accounts for agreement between the computer predictions and human codes by chance, while MCC is a metric that only provides a high score if the model performs well for all four quadrants of the confusion matrix—that is, true negatives (TN), false negatives (FN), false positives (FP), and true positives (TP).[58] The confusion matrices in Table 9.2 show the numbers of responses predicted by the computer in comparison to how the responses were human-coded and indicates the number of TN, FN, FP, and TP as predicted by the models. The values presented for the accuracy, Cohen's kappa, and MCC metrics are the mean and standard deviations after iterating each model 30 times; the confusion matrix indicates the averages of the

30 iterations. As shown in Table 9.2, the accuracy values for all models, except for the cause–effect only model, are above 87% with Cohen's kappa and MCC values exceeding 0.7 (i.e., the standard cut-off for using ATA models in assessment).[64] These results indicate that the CNN models can predict the presence of each feature across students' written responses that describe and explain the three different reaction types, with the models for several features reaching near-perfect agreement with human coding (i.e., reaction medium, charges, and electron movement all have kappa and MCC > 0.95). Models for connectivity of molecules and stereochemistry features that both have dynamic counterparts (i.e., bond breaking and making and formation of stereochemistry, respectively) tended to perform less well than models for static features without a similar dynamic feature. In this case, multiple features have similar markers (i.e., words and phrases) in students' writing, making the models more difficult to train. The cause–effect only model reached an accuracy of 79% with Cohen's kappa and MCC of 0.585. Although this model did not perform as well as the other models, it still performs quite well for an ATA model of cause-and-effect reasoning, particularly with the use of a CNN model, which requires minimal computational resources to train. Cause and effect reasoning is challenging for an ATA model to recognize because it can be nuanced or implied within the writing (i.e., markers of cause and effect such as "because of" or "due to" may not always be present).[65–67] Typically, complex machine learning models, which require much greater computational resources to train, must be used to model more complex reasoning within the text.[9,66,68,69] We believe that our electronic causal reasoning model performed better than the cause–effect only model due to a greater presence of explicit markers in students' writing, such as the indication of electronic properties alongside language about electron movement.

Table 9.2 Description of the model for each feature of mechanistic reasoning including the total number of sentences used to develop the model, the size of the testing set, accuracy, Cohen's kappa, and Matthews correlation coefficient (MCC). A confusion matrix for each is also included. TN = true negative, FN = false negative, FP = false positive, TP = true positive. All statistics represent average scores across 30 models; values in parentheses are the standard deviations

Feature	***N* (total sentences)**	***n* (testing set)**	**Accuracy**	**Cohen's kappa**	**MCC**	**Confusion matrix**	
						TN	**FN**
						FP	**TP**

Reaction medium	1343	135	0.990 (0.004)	0.958 (0.016)	0.959 (0.015)	116	0
						1.333	17.667
Connectivity of molecules	1611	162	0.884 (0.011)	0.768 (0.023)	0.769 (0.023)	70.233	9.767
						9	73
Charges	1243	125	0.989 (0.006)	0.975 (0.014)	0.975 (0.013)	86	0
						1.333	37.667
Stereochemistry	1243	125	0.914 (0.009)	0.741 (0.027)	0.749 (0.028)	93.733	2.267
						8.433	20.567
Electron movement	1243	125	0.997 (0.004)	0.993 (0.010)	0.993 (0.010)	89.633	0.367
						0	35
Non-electronic mechanism	1762	173	0.871 (0.015)	0.738 (0.030)	0.739 (0.030)	63.867	13.133
						9.167	86.833
Bond breaking and making	1243	125	0.952 (0.007)	0.825 (0.023)	0.826 (0.023)	101.3	2.7
						3.333	17.667
Implicit properties	1243	125	0.943 (0.007)	0.792 (0.027)	0.792 (0.028)	100.8	4.2
						2.967	17.033
Formation of stereochemistry	1458	146	0.942 (0.011)	0.819 (0.033)	0.819 (0.033)	112.433	4.567
						3.933	25.067
Cause–effect only	1976	198	0.793 (0.011)	0.585 (0.023)	0.585 (0.022)	72.267	20.733
						20.2	84.8

Electronic causal	1534	155	0.939 (0.008)	0.835 (0.021)	0.836 (0.021)	112.533	3.467
						5.967	33.033

9.8 Implications

9.8.1 Implications for research

We have presented a modified version of the mechanistic reasoning framework presented by Russ et al.[51] and adapted by Watts et al.[8] as a lens through which to analyze students' mechanistic reasoning about three different organic chemistry reaction mechanisms. Other researchers can use this modified framework to study how students reason about additional reaction mechanisms and to study how specific writing prompts elicit features of mechanistic reasoning. Researchers can use similar methods to develop ATA models to evaluate students' writing using other frameworks and writing prompts. These automated models can be used for further research to develop instructor- and student-facing feedback platforms that can provide formative, immediate feedback. Furthermore, automated models can be used to evaluate the impact of interventions on students' use of features of mechanistic reasoning, such as tutorials based on findings in the literature.[5,6]

9.8.2 Implications for practice

Instructors can use or modify the WTL prompts described in this study to support students' mechanistic reasoning. Furthermore, our findings indicate the careful attention instructors should place when developing prompts and ensuring that the language is specific to elicit the desired responses from students. Instructors can also use the framework described in this study to develop constructed-response items and WTL prompts that elicit features of mechanistic reasoning. Instructors may use these items in their courses for both formative and summative assessment and can use the framework to support students' considerations of all components necessary for mechanistic reasoning. If you are an instructor and are interested in using our ATA models for formative assessment, please contact us.

9.9 Limitations

The work presented in this chapter provides an example of a broadly applicable mechanistic reasoning framework and corresponding ATA models which have been used for students' written descriptions and explanations of three different mechanisms. However, there are many reaction mechanisms in the organic chemistry curriculum, and we cannot assume that the framework and models will perform similarly with other mechanisms. Further testing with different mechanism types is required to broaden the scope of these models. Additionally, all the training and testing data was collected from students' final drafts in one course at one university. There are many ways to describe and explain mechanisms that the students in this study may not have exemplified in their final drafts. More work is required with additional data sources and populations of students to determine if the models perform similarly with different populations of students. Lastly, the approach of analyzing students' writing is limited in that their writing might not capture their full understanding or reasoning. While the models presented here are a useful tool to provide students and instructors with formative feedback, we do not recommend their use for summative feedback due to the limited population the data came from.

9.10 Conclusions

This study presented a modified version of Russ et al. and Watts et al.'s mechanistic reasoning framework applied to students' written explanations of three different organic chemistry reaction mechanisms. The presence of features necessary for mechanistic reasoning varied based on the nature of the writing prompt and the specific mechanism students were explaining. Students' responses were used to train several ATA models that can successfully predict whether students included these features in their written responses. The results of this study show that the modified mechanistic reasoning framework presented can be applied to identify features of students' writing across multiple types of mechanisms. The analysis across students' writing for different mechanisms indicate that the nature of the prompt influences how mechanistic reasoning is elicited. This study additionally indicates that CNN models can be used to successfully provide automated identification of features necessary for mechanistic reasoning in students' writing. These findings extend the literature by indicating that ATA models can be successfully deployed across students' writing for up to three different organic reaction mechanisms with utility for identifying students' written descriptions of *how* and *why* mechanisms occur.

9.11 Appendix

Thalidomide WTL assignment

Developing a Therapeutic Analog for Thalidomide

Thalidomide was widely used after World War II as a sedative and later as a treatment for morning sickness. Unfortunately, after its widespread use, it was discovered that thalidomide causes very serious side effects—in particular, birth defects such as phocomelia (limb malformation). The drug was banned in 1962, and these events resulted in important changes to the way the FDA approves drugs. Now, despite the inherent dangers, thalidomide is used for treatment of nausea related to chemotherapy, where benefit of treatment outweighs the inherent dangers.

It is understood that thalidomide exists as two **enantiomers**; one is a teratogen that causes birth defects, while the other has therapeutic properties. Rapid **racemization** occurs at neutral pH, so both enantiomers are formed at roughly an equal mixture in the blood, which means that, even if only the therapeutic isomer is used, both will form once introduced in the body. The racemization is illustrated below in Figure 9.2.

Figure 9.2 The rapid racemization of thalidomide.

Furthermore, both enantiomers are subject to **acid hydrolysis** once in the stomach at lower pH, which could produce products that are teratogens. The structure of thalidomide and two thalidomide hydrolysis products are shown below in Figure 9.3. For these reasons, it is important to prevent both the racemization and the subsequent hydrolysis of thalidomide.

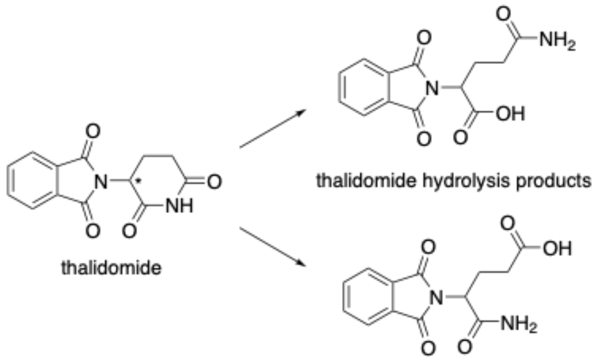


Figure 9.3 Thalidomide and two thalidomide hydrolysis products. The stereocenter is shown (*).

You are an OB-GYN at the Mayo Clinic. A colleague, who is an oncologist at the University of Minnesota, has approached you about a potential collaboration on a human clinical trial. This trial will propose and test the efficacy of thalidomide **analogs** for the treatment of nausea in cancer patients. (See note on the third page for an explanation of an analog.)

As an organic expert in the chemical pathways that lead to birth defects, you are writing an email to your collaborator. Your goal will be to propose a structural difference that will make the thalidomide analog unreactive toward both racemization and hydrolysis. You must provide descriptions of the structure and reactivity of thalidomide toward racemization and hydrolysis as well as descriptions of the structural differences in the proposed analog that will make it unreactive to both of these processes. The oncologist is not an expert in organic chemistry. Therefore, carefully consider which organic chemistry terms to use and when to define or explain them. Use clear and concise language, striking a balance between organic jargon and oversimplified explanations.

Your email should be approximately between 500-700 words (1-2 pages) in length. It should address the following points:

1. Provide thorough descriptions of the mechanisms of both racemization and acid hydrolysis, highlighting the critical structural features of thalidomide and their role in these mechanisms.
    a. When racemization occurs, what changes occur in the molecule?
    b. When hydrolysis occurs, what changes occur in the molecule?

2. Propose a thalidomide analog (one compound) that would not undergo racemization or hydrolysis. Explain what structural features are in place that would inhibit or prevent these processes.

You can and should include figures of schemes, structures, or mechanisms, if that supports your response. We suggest that you have the figure(s) in front of you—ready to color-code or mark-up in various ways—and that you use your visible thinking to guide your audience through your explanation. Any images that you include in your response, *including the figures in this prompt or those that you draw in ChemDraw or on paper*, must have the original source cited using either ACS or APA format. Given your audience, your written response should suffice so that the explanations can be understood without the figures. **You will be graded only on your written response.**

An analog is a compound that is very similar to but has small structural differences from the pharmaceutical target. For example, *m*-cresol (shown in Figure 9.4 below) is an analog of phenol.

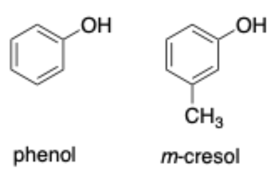


Figure 9.4 Phenol and m-cresol, an analog of phenol.

Thalidomide WTL assignment peer review rubric

Peer Review Guidelines:

- Print and read over your peer's essay to quickly get an overview of the piece.
- Read the essay more slowly keeping the rubric in mind.
- Highlight the pieces of texts that let you directly address the rubric prompts in your online responses.
- In your online responses, focus on larger issues (higher order concerns) of content and argument rather than lower order concerns like grammar and spelling.
- Be very specific in your responses, referring to your peer's actual language, mentioning terms and concepts that are either present or missing, and following the directions in the rubric.

- Use respectful language whether you are suggesting improvements to or praising your peer.

1. How well does the author explain the process of racemization in thalidomide? Suggest some ways that the author could improve their mechanism description, including discussing what changes occur in the thalidomide molecule through the racemization mechanism.
2. How well does the author explain the process of hydrolysis in thalidomide? Suggest some ways that the author could improve their mechanism description, including discussing what changes occur in the thalidomide molecule through the hydrolysis mechanism.
3. Does the author propose a reasonable thalidomide analog that would not undergo racemization or hydrolysis? To what extent does the author explain the specific structural features that are present in the thalidomide analog that would stop racemization and/or hydrolysis from occurring?
4. In what ways did the author use and define organic chemistry terms that may be unfamiliar to a non-organic chemist? Comment on ways the author could enhance the clarity of their email by translating the definitions of the organic chemistry into their own words.

Wittig WTL assignment

Using the Base-Free Wittig Reaction to Synthesize Anticancer Compounds

Benzoxepine (Figure 9.5) is a heterocycle composed of a six-membered benzene ring and a seven-membered oxepin ring. Some benzoxepine analogs inhibit tuberculosis, and others inhibit cancers by inducing activation of the apoptosis pathway. The benzoxepine analog shown in Figure 9.6 is a **benzoxepinoisoxazolone** whose anticancer activity is attributed to its structure that is functualized with phenyl and azole groups.

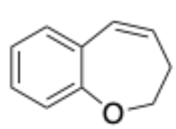


Figure 9.5 Benzoxepine.

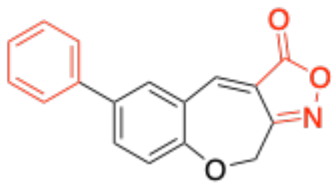


Figure 9.6 A benzoxepinoisoxazolone, a benzoxepine that has been modified with phenyl and azole functional groups.

However useful, isolating benzoxepine analogs from natural sources is inefficient. Benzoxepine analogs are important intermediates in the synthesis of therapeutic drugs, such as the aforementioned benzoxepinoisoxazolone. They are also important in studies that deduce structure-activity relationships to develop other medicinal treatments. Recently, German researchers synthesized benzoxepine analogs (Figure 9.7) using a **base-free Wittig reaction** (Figure 9.8). This reaction is a novel development that will synthesize therapeutic drugs on an industrial scale while producing fewer waste byproducts.

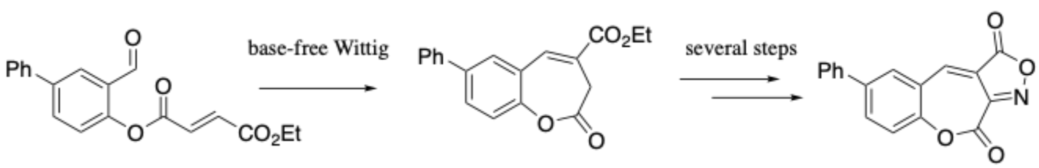


Figure 9.7 Synthesis of benzoxepinoisoxazolone through the base-free Wittig reaction.

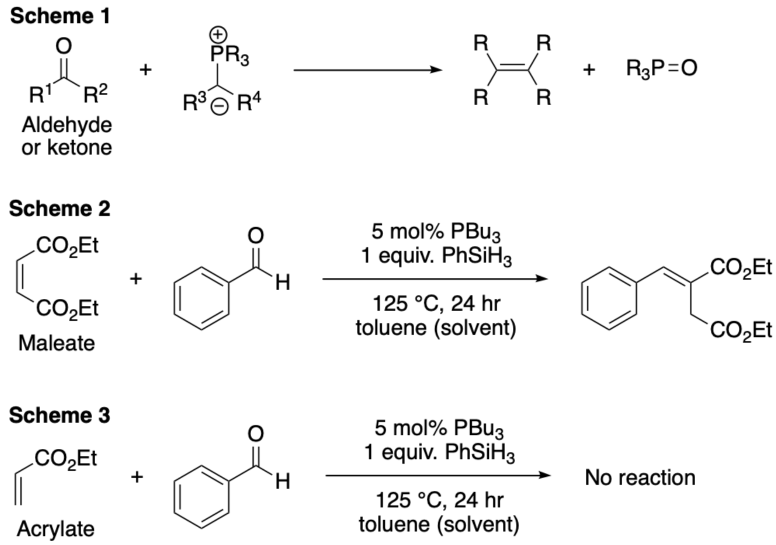


Figure 9.8 Generalized schemes of the base-free Wittig reaction. Scheme 1 shows the standard Wittig reaction, and Scheme 2 shows an example of the base-free Wittig reaction using a maleate starting material. Scheme 3 shows that the base-free Wittig reaction fails when using an acrylate starting material instead.

You are a medicinal drug developer in a research group that primarily studies anticancer compounds. Inspired by the benzoxepinoisoxazolone in Figure 9.6, the group's current goal is synthesizing benzoxepine analogs using the already developed base-free Wittig synthesis and evaluating them for anticancer activities. To do so, your research team is drafting a grant proposal for the National Institute of Health (NIH) that summarizes the group's research goals and argues for the significance, innovation, and impact. You, the organic chemist expert, must write the section of the grant proposal that explains the base-free Wittig reaction that synthesizes benzoxepine analogs. Because the reaction is critical for the success of the project, you must demonstrate to the committee that your team understands how the reaction works and why it is selective. The committee who will review the proposal is made up of scientists from many disciplines, including chemistry, biology, and medicine. Therefore, they may not be experts when it concerns mechanisms or organic-specific terms. The NIH recommends that you:

- write organized and logical paragraphs
- include figures that assist the reviewers in understanding complex information
- use clear and concise language, striking a balance between organic jargon and oversimplified explanations

Your section of the grant proposal should be approximately between 500-700 words (1-2 pages) in length. It should address the following points:

1. Explain the critical structural and electronic features and properties of the starting materials and reagents in Scheme 2 and their role in the mechanistic steps that lead to the formation of the products without the use of an external base.
    a. In describing the mechanistic steps for the reaction in Scheme 2, what changes occur within those steps to the starting materials and reagents that lead to the formation of the ylide? (Note that the ylide is not shown in this scheme.)
    b. What structural changes happen to PBu_3 at each mechanistic step?
    c. Focus on the *how* and *why* as well as the *what.*
2. When comparing the starting materials and reagents in Scheme 2 to those in Scheme 1, what structural differences are present that allow the Wittig reaction to proceed without the use of an external base?
3. Why would researchers want to synthesize benzoxepinones through the modified, base-free Wittig reaction over the traditional Wittig reaction? Focus on key aspects of the overall reaction that make it significant, innovative, and impactful for larger-scale research studies.
4. Propose a reason why the reaction works with maleate but does not work with acrylate, as shown in Scheme 3. What structural features are present or absent in the acrylate that prevent the modified Wittig mechanism from happening?

You can and should include figures of schemes, structures, or mechanisms, if that supports your response. We suggest that you have the figure(s) in front of you—ready to color-code or mark-up in various ways—and that you use your visible thinking to guide your audience through your explanation. Any images that you include in your response, *including the figures in this prompt or those that you draw in ChemDraw or on paper*, must have the original source cited using either ACS or APA format. Given your audience, your written response should suffice so that <u>the explanations can be understood without the figures</u>. **You will be graded only on your written response.**

Wittig WTL assignment peer review rubric

Peer Review Guidelines:

- Print and read over your peer's essay to quickly get an overview of the piece.

- Read the essay more slowly keeping the rubric in mind.
- Highlight the pieces of texts that let you directly address the rubric prompts in your online responses.
- In your online responses, focus on larger issues (higher order concerns) of content and argument rather than lower order concerns like grammar and spelling.
- Be very specific in your responses, referring to your peer's actual language, mentioning terms and concepts that are either present or missing, and following the directions in the rubric.
- Use respectful language whether you are suggesting improvements to or praising your peer.

1. To what extent does the author explain each mechanistic step in Scheme 2? Suggest some ways that the author could improve their mechanistic description, including discussing what changes occur to the starting materials, reagents, and PBu_3 that lead to the formation of the ylide.
2. Does the author compare the structural differences in the starting materials and reagents between Scheme 1 and Scheme 2? To what extent do their descriptions explain why the reaction in Scheme 2 does not need an external base, which makes it attractive for industrial use?
3. Does the author provide an explanation for why researchers would want to synthesize benzoxepinones through the modified, base-free Wittig reaction? Comment on ways the author can improve their explanations by focusing on key aspects of the overall reaction that make it significant for large-scale applications.
4. Does the author explain why the reaction conditions lead to no reaction happening in Scheme 3? To what extent does the author describe the specific structural features of acrylate that prevent the modified Wittig reaction from occurring?

9.12 References

(1) Cooper, M.; Kouyoumdjian, H.; Underwood, S. Investigating Students' Reasoning about Acid–Base Reactions. *J. Chem. Educ.* **2016**, *93* (10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417.

(2) Crandell, O.; Kouyoumdjian, H.; Underwood, S.; Cooper, M. Reasoning about Reactions in Organic Chemistry: Starting It in General Chemistry. *J. Chem. Educ.* **2018**, *96* (2), 213–226. https://doi.org/10.1021/acs.jchemed.8b00784.

(3) Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. Arrows on the Page Are Not a Good Gauge: Evidence for the Importance of Causal Mechanistic Explanations about Nucleophilic Substitution in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (2), 313–327. https://doi.org/10.1021/acs.jchemed.9b00815.

(4) Dood, A. J.; Dood, J. C.; Cruz-Ramírez de Arellano, D.; Fields, K. B.; Raker, J. R. Using the Research Literature to Develop an Adaptive Intervention to Improve Student Explanations of an SN1 Reaction Mechanism. *J. Chem. Educ.* **2020**, *97* (10), 3551–3562. https://doi.org/10.1021/acs.jchemed.0c00569.

(5) Dood, A. J.; Fields, K. B.; Raker, J. R. Using Lexical Analysis To Predict Lewis Acid–Base Model Use in Responses to an Acid–Base Proton-Transfer Reaction. *J. Chem. Educ.* **2018**, *95* (8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177.

(6) Dood, A. J.; Dood, J. C.; Cruz-Ramírez de Arellano, D.; Fields, K. B.; Raker, J. R. Analyzing Explanations of Substitution Reactions Using Lexical Analysis and Logistic Regression Techniques. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 267–286. https://doi.org/10.1039/C9RP00148D.

(7) Dood, A. J.; Fields, K. B.; Cruz-Ramírez de Arellano, D.; Raker, J. R. Development and Evaluation of a Lewis Acid–Base Tutorial for Use in Postsecondary Organic Chemistry Courses. *Can. J. Chem.* **2019**, *97* (10), 711–721. https://doi.org/10.1139/cjc-2018-0479.

(8) Watts, F. M.; Schmidt-McCormack, J. A.; Wilhelm, C. A.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21* (4), 1148–1172. https://doi.org/10.1039/C9RP00185A.

(9) Winograd, B. A.; Dood, A. J.; Moeller, R.; Moon, A.; Gere, A.; Shultz, G. Detecting High Orders of Cognitive Complexity in Students' Reasoning in Argumentative Writing About Ocean Acidification. In *LAK21: 11th International Learning Analytics and Knowledge Conference*; LAK21; Association for Computing Machinery: New York, NY, USA, 2021; pp 586–591. https://doi.org/10.1145/3448139.3448202.

(10) Yik, B. J.; Dood, A. J.; Arellano, D. C.-R. de; Fields, K. B.; Raker, J. R. Development of a Machine Learning-Based Tool to Evaluate Correct Lewis Acid–Base Model Use in Written Responses to Open-Ended Formative Assessment Items. *Chem. Educ. Res. Pract.* **2021**, *22*, 866–885. https://doi.org/10.1039/D1RP00111F.

(11) Noyes, K.; McKay, R. L.; Neumann, M.; Haudek, K. C.; Cooper, M. M. Developing Computer Resources to Automate Analysis of Students' Explanations of London Dispersion Forces. *J. Chem. Educ.* **2020**, *97* (11), 3923–3936. https://doi.org/10.1021/acs.jchemed.0c00445.

(12) Winograd, B. A.; Dood, A. J.; Finkenstaedt-Quinn, S. A.; Gere, A. R.; Shultz, G. V. Automating Characterization of Peer Review Comments in Chemistry Courses. In *14th Computer-Supported Collaborative Learning (CSCL) - Proceedings*; Bochum, Germany, 2021; pp 11–18.

(13) Goodwin, W. Mechanisms and Chemical Reaction. In *Philosophy of Chemistry*; Woody, A. I., Hendry, R. F., Needham, P., Eds.; Handbook of the Philosophy of Science; North-Holland: Amsterdam, 2012; Vol. 6, pp 309–327. https://doi.org/10.1016/B978-0-444-51675-6.50023-2.

(14) Goodwin, W. Explanation in Organic Chemistry. *Annals of the New York Academy of Sciences* **2003**, *988* (1), 141–153. https://doi.org/10.1111/j.1749-6632.2003.tb06093.x.

(15) Kozma, R. The Material Features of Multiple Representations and Their Cognitive and Social Affordances for Science Understanding. *Learning and Instruction* **2003**, *13* (2), 205–226.

(16) Bhattacharyya, G. From Source to Sink: Mechanistic Reasoning Using the Electron-Pushing Formalism. *J. Chem. Educ.* **2013**, *90* (10), 1282–1289. https://doi.org/10.1021/ed300765k.

(17) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402–1407. https://doi.org/10.1021/ed082p1402.

(18) Grove, N.; Cooper, M.; Rush, K. Decorating with Arrows: Toward the Development of Representational Competence in Organic Chemistry. *J. Chem. Educ.* **2012**, *89* (7), 844–849. https://doi.org/10.1021/ed2003934.

(19) Ferguson, R.; Bodner, G. M. Making Sense of the Arrow-Pushing Formalism among Chemistry Majors Enrolled in Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 102–113. https://doi.org/10.1039/B806225K.

(20) 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.

(21) Anderson, T. L.; Bodner, G. M. What Can We Do about 'Parker'? A Case Study of a Good Student Who Didn't 'Get' Organic Chemistry. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 93–101. https://doi.org/10.1039/B806223B.

(22) Wilson, S. B.; Varma-Nelson, P. Characterization of First-Semester Organic Chemistry Peer-Led Team Learning and Cyber Peer-Led Team Learning Students' Use and Explanation of Electron-Pushing Formalism. *J. Chem. Educ.* **2019**, *96* (1), 25–34. https://doi.org/10.1021/acs.jchemed.8b00387.

(23) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Ideas about Nucleophiles and Electrophiles: The Role of Charges and Mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797–810. https://doi.org/10.1039/C5RP00113G.

(24) Anzovino, M. E.; Bretz, S. L. Organic Chemistry Students' Fragmented Ideas about the Structure and Function of Nucleophiles and Electrophiles: A Concept Map Analysis. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 1019–1029. https://doi.org/10.1039/C6RP00111D.

(25) DeFever, R. S.; Bruce, H.; Bhattacharyya, G. Mental Rolodexing: Senior Chemistry Majors' Understanding of Chemical and Physical Properties. *J. Chem. Educ.* **2015**, *92* (3), 415–426. https://doi.org/10.1021/ed500360g.

(26) Cooper, M.; Grove, N.; Underwood, S.; Klymkowsky, M. Lost in Lewis Structures: An Investigation of Student Difficulties in Developing Representational Competence. *J. Chem. Educ.* **2010**, *87* (8), 869–874. https://doi.org/10.1021/ed900004y.

(27) Cooper, M.; Corley, L.; Underwood, S. An Investigation of College Chemistry Students' Understanding of Structure–Property Relationships. *J Res Sci Teach* **2013**, *50* (6), 699–721. https://doi.org/10.1002/tea.21093.

(28) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable Reasoning: Multi-Variate Problem-Solving in Organic Chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281–292. https://doi.org/10.1039/C0RP90003F.

(29) 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.

(30) Bhattacharyya, G. Who Am I? What Am I Doing Here? Professional Identity and the Epistemic Development of Organic Chemists. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 84–92. https://doi.org/10.1039/B806222F.

(31) Finkenstaedt-Quinn, S. A.; Watts, F. M.; Petterson, M. N.; Archer, S. R.; Snyder-White, E. P.; Shultz, G. V. Exploring Student Thinking about Addition Reactions. *J. Chem. Educ.* **2020**, *97* (7), 1852–1862. https://doi.org/10.1021/acs.jchemed.0c00141.

(32) Petterson, M. N.; Watts, F. M.; Snyder-White, E. P.; Archer, S. R.; Shultz, G. V.; Finkenstaedt-Quinn, S. A. Eliciting Student Thinking about Acid–Base Reactions via App and Paper–Pencil Based Problem Solving. *Chem. Educ. Res. Pract.* **2020**, *21* (3), 878–892. https://doi.org/10.1039/C9RP00260J.

(33) Watts, F. M.; Zaimi, I.; Kranz, D.; Graulich, N.; Shultz, G. V. Investigating Students' Reasoning over Time for Case Comparisons of Acyl Transfer Reaction Mechanisms. *Chem. Educ. Res. Pract.* **2021**, *22* (2), 364–381. https://doi.org/10.1039/D0RP00298D.

(34) Cruz-Ramírez de Arellano, D.; Towns, M. H. Students' Understanding of Alkyl Halide Reactions in Undergraduate Organic Chemistry. *Chemistry Education Research and Practice* **2014**, *15* (4), 501–515. https://doi.org/10.1039/C3RP00089C.

(35) Graulich, N.; Hedtrich, S.; Harzenetter, R. Explicit versus Implicit Similarity – Exploring Relational Conceptual Understanding in Organic Chemistry. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 924–936. https://doi.org/10.1039/C9RP00054B.

(36) 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.

(37) Caspari, I.; Weinrich, M. L.; Sevian, H.; Graulich, N. This Mechanistic Step Is "Productive": Organic Chemistry Students' Backward-Oriented Reasoning. *Chem. Educ. Res. Pract.* **2018**, *19* (1), 42–59. https://doi.org/10.1039/C7RP00124J.

(38) Caspari, I.; Kranz, D.; Graulich, N. Resolving the Complexity of Organic Chemistry Students' Reasoning through the Lens of a Mechanistic Framework. *Chem. Educ. Res. Pract.* **2018**, *19*, 1117–1141. https://doi.org/10.1039/C8RP00131F.

(39) Cooper, M.; Klymkowsky, M. Chemistry, Life, the Universe, and Everything: A New Approach to General Chemistry, and a Model for Curriculum Reform. *J. Chem. Educ.* **2013**, *90* (9), 1116–1122. https://doi.org/10.1021/ed300456y.

(40) Cooper, M.; Stowe, R. L.; Crandell, O. M.; Klymkowsky, M. W. Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* **2019**, *96* (9), 1858–1872. https://doi.org/10.1021/acs.jchemed.9b00401.

(41) Moon, A.; Moeller, R.; Gere, A. R.; Shultz, G. V. Application and Testing of a Framework for Characterizing the Quality of Scientific Reasoning in Chemistry Students' Writing on Ocean Acidification. *Chem. Educ. Res. Pract.* **2019**, *20* (3), 484–494. https://doi.org/10.1039/C9RP00005D.

(42) Grimberg, B. I.; Hand, B. Cognitive Pathways: Analysis of Students' Written Texts for Science Understanding. *International Journal of Science Education* **2009**, *31* (4), 503–521. https://doi.org/10.1080/09500690701704805.

(43) Moon, A.; Gere, A. R.; Shultz, G. V. Writing in the STEM Classroom: Faculty Conceptions of Writing and Its Role in the Undergraduate Classroom. *Science Education* **2018**, *102* (5), 1007–1028. https://doi.org/10.1002/sce.21454.

(44) Flower, L.; Hayes, J. R. A Cognitive Process Theory of Writing. *College Composition and Communication* **1981**, *32* (4), 365–387. https://doi.org/10.2307/356600.

(45) Flower, L.; Hayes, J. R. Images, Plans, and Prose: The Representation of Meaning in Writing. *Written Communication* **1984**, *1* (1), 120–160.

https://doi.org/10.1177/0741088384001001006.

(46) Hayes, J. R. A New Framework for Understanding Cognition and Affect in Writing. In *The Science of Writing*; Levy, C. M., Ransdell, S., Eds.; Routledge, 1996.

(47) Finkenstaedt-Quinn, S. A.; Petterson, M.; Gere, A.; Shultz, G. Praxis of Writing-to-Learn: A Model for the Design and Propagation of Writing-to-Learn in STEM. *J. Chem. Educ.* **2021**, *98* (5), 1548–1555. https://doi.org/10.1021/acs.jchemed.0c01482.

(48) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Research in the Teaching of English* **2015**, *50* (2), 199–235.

(49) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and Conceptual Learning in Science: An Analysis of Assignments. *Written Communication* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(50) Gupte, T.; Watts, F. M.; Schmidt-McCormack, J. A.; Zaimi, I.; Gere, A. R.; Shultz, G. V. Students' Meaningful Learning Experiences from Participating in Organic Chemistry Writing-to-Learn Activities. *Chem. Educ. Res. Pract.* **2021**, *22* (2), 396–414. https://doi.org/10.1039/D0RP00266F.

(51) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Science Education* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(52) Machamer, P. K.; Darden, L.; Craver, C. F. Thinking About Mechanisms. *Philosophy of Science* **2000**, *67* (1), 1–25. https://doi.org/10.1086/392759.

(53) Darden, L. Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/Backward Chaining. *Philosophy of Science* **2002**, *69* (S3), S354–S365. https://doi.org/10.1086/341858.

(54) Keiner, L.; Graulich, N. Transitions between Representational Levels: Characterization of Organic Chemistry Students' Mechanistic Features When Reasoning about Laboratory Work-up Procedures. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 469–482. https://doi.org/10.1039/C9RP00241C.

(55) Kirilenko, A. P.; Stepchenkova, S. Inter-Coder Agreement in One-to-Many Classification: Fuzzy Kappa. *PLOS ONE* **2016**, *11* (3), e0149787. https://doi.org/10.1371/journal.pone.0149787.

(56) Watts, F. M.; Finkenstaedt-Quinn, S. A. The Current State of Methods for Establishing Reliability in Qualitative Chemistry Education Research Articles. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 565–578. https://doi.org/10.1039/D1RP00007A.

(57) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825−2830.

(58) van Rossum, G.; Drake, F. L. *The Python Language Reference Manual*; Network Theory Ltd., 2011.

(59) Chollet, F.; others. Keras. 2015.

(60) Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **1960**, *20*, 37–46. https://doi.org/10.1177/001316446002000104.

(61) Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genomics* **2020**, *21* (1), 6. https://doi.org/10.1186/s12864-019-6413-7.

(62) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521* (7553), 436–444. https://doi.org/10.1038/nature14539.

(63) Chollet, F. Deep Learning for Text and Sequences. In *Deep Learning with Python*; Manning Publications Company, 2018; pp 178–232.

(64) Williamson, D. M.; Xi, X.; Breyer, F. J. A Framework for Evaluation and Use of Automated Scoring. *Educational Measurement: Issues and Practice* **2012**, *31* (1), 2–13. https://doi.org/10.1111/j.1745-3992.2011.00223.x.

(65) Blanco, E.; Castell, N.; Moldovan, D. Causal Relation Extraction. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*; European Language Resources Association (ELRA): Marrakech, Morocco, 2008.

(66) Dasgupta, T.; Saha, R.; Dey, L.; Naskar, A. Automatic Extraction of Causal Relations from Text Using Linguistically Informed Deep Neural Networks. In *SIGDIAL Conference*; 2018.

(67) Asghar, N. Automatic Extraction of Causal Relations from Natural Language Texts: A Comprehensive Survey. *arXiv:1605.07895 [cs]* **2016**.

(68) Hendrickx, I.; Kim, S. N.; Kozareva, Z.; Nakov, P.; Séaghdha, D. Ó.; Padó, S.; Pennacchiotti, M.; Romano, L.; Szpakowicz, S. SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations Between Pairs of Nominals. *arXiv:1911.10422 [cs]* **2019**.

(69) Zhao, S.; Wang, Q.; Massung, S.; Qin, B.; Liu, T.; Wang, B.; Zhai, C. Constructing and Embedding Abstract Event Causality Networks from Text Snippets. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*; WSDM '17;

Association for Computing Machinery: New York, NY, USA, 2017; pp 335–344. https://doi.org/10.1145/3018661.3018707.

Chapter 10

Towards Developing an Interactive Tool To Provide Automated, Formative Feedback on Students' Written Descriptions of Organic Reaction Mechanisms

10.1 Initial remarks

This chapter presents the development of a tool for delivering automated, formative feedback to students' responses to the writing-to-learn (WTL) assignments which elicit students' mechanistic reasoning. The research in this chapter leverages the machine learning models described in Chapter 9 to automatically analyze student writing and present students with tailored formative feedback. The chapter specifically describes the design and development of the automated feedback tool, alongside presenting findings from pilot interviews with students using the tool. This research contributes to the literature by demonstrating how machine learning models for identifying features of students' mechanistic reasoning can be used to build tools for providing automated feedback. Furthermore, this study provides initial evidence suggesting how students might use automated feedback when responding to WTL assignments.

The study describes the development of the automated feedback tool with the goal to promote students' self-regulated learning, which is a framework for understanding how self-regulated processes (such as cognition and behavior) mediate how students set goals and develop strategies for achieving internally set learning outcomes for a given task. The feedback tool design is described in alignment with a framework for classifying automated feedback technologies, and qualitative methodologies were used to evaluate the tool in pilot interviews with students from the second-semester organic chemistry laboratory course. The methods for designing the tool detail the considerations regarding how students are intended to use the feedback tool during the WTL assignment process, how the tool processes student data, and how the tool generates feedback based on individual student responses. The individualized feedback is aligned with the machine learning models and based on the literature regarding students' learning of mechanistic reasoning in organic chemistry. The results of the study present the feedback tool itself, which provides a textbox for students to submit their writing, produces a graphical display of the results of the automated analysis, and provides tailored, written feedback to guide students' revisions. The interview analysis suggests how the feedback tool can further communicate the expectations for students' responses beyond the assignment rubric, promote students' metacognitive processes, and influence how students engage in the WTL structures in place to support interactive writing (i.e., peer review and writing fellows). For example, students discussed how the feedback would support their reflection on what is important to include in their response to the assignment and help them identify areas to revise. The students also indicated that using the tool would complement, rather than replace, the way they use peer review and writing fellows during the writing process. The findings from this study provide initial evidence that students' use of the automated feedback tool would align with and support the effective components of WTL pedagogy, such as engaging in metacognition and participating in opportunities for interactive writing.

This chapter will be published in the forthcoming peer-reviewed conference proceedings *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13—17, 2023, Arlington, TX, USA*. The original publication and copyright information are provided below. The publication was modified to adhere to Rackham dissertation formatting requirements, and no additional changes were made. As primary author, I contributed to conceptualization, methodology, data collection, analysis, and writing (both original draft preparation and review and editing). A.J. Dood contributed to conceptualization, methodology, data collection, analysis, and writing (review and editing). G.V. Shultz contributed to project supervision, conceptualization, and writing (review and editing.)

<u>*Original publication and copyright information:*</u>

Reproduced from F.M. Watts, A.J. Dood, and G.V. Shultz, in *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13—17, 2023, Arlington, TX, USA*, Association for Computing Machinery, 2023, in press. Copyright 2023 ACM.

10.2 Abstract

Writing-to-learn (WTL) pedagogy supports the implementation of writing assignments in STEM courses to engage students in conceptual learning. Recent studies in the undergraduate STEM context demonstrate the value of implementing WTL, with findings that WTL can support meaningful learning and elicit students' reasoning. However, the need for instructors to provide

feedback on students' writing poses a significant barrier to implementing WTL; this barrier is especially notable in the context of introductory organic chemistry courses at large universities, which often have large enrollments. This work describes one approach to overcome this barrier by presenting the development of an automated feedback tool for providing students with formative feedback on their responses to an organic chemistry WTL assignment. This approach leverages machine learning models to identify features of students' mechanistic reasoning in response to WTL assignments in a second-semester, introductory organic chemistry laboratory course. The automated feedback tool development was guided by a framework for designing automated feedback, theories of self-regulated learning, and the components of effective WTL pedagogy. Herein, we describe the design of the automated feedback tool and report our initial evaluation of the tool through pilot interviews with organic chemistry students.

10.3 Introduction

Writing-to-learn (WTL) pedagogy involves implementing writing assignments to support students' conceptual learning.[1,2] Within the context of organic chemistry courses, studies provide evidence that WTL effectively engages students in articulating their thinking, constructing new ideas, and building connections between content.[3–6] Furthermore, WTL promotes meaningful learning by contributing to affective components of the learning process, such as students' motivation and interest.[7,8] Despite the evidence that WTL is an effective instructional practice, instructors may not implement WTL in their courses because of the time required to provide students with feedback. The ability to provide feedback is especially a concern for large-enrollment courses.[9,10] As such, there is a need for research to overcome this barrier to the implementation of WTL pedagogy. This work presents an approach to this problem by leveraging machine learning (ML) to automatically analyze students' writing and generate content-focused, formative feedback.

10.4 Background

WTL promotes the incorporation of writing assignments to engage students in articulating their thinking and expanding their conceptual understanding.[1,2] WTL is most effective when assignments involve meaning-making tasks that require students to move beyond restating conceptual knowledge, such as using their conceptual understanding to evaluate data or construct arguments. Furthermore, effective WTL assignments should be implemented with clear expectations, opportunities for interactive writing, and structures that promote metacognition.[1] Setting clear expectations involves indicating what students are supposed to do and how they will be evaluated; assignment descriptions and rubrics must reflect the expectations for students' learning. Opportunities for interactive writing include any interactions that occur before the final draft of the assignment is due, such as peer review or discussing the assignment with writing tutors. Lastly, promoting metacognition involves encouraging students to reflect on their understanding as they complete the assignment. This study is in the context of a WTL implementation with required peer review and revision to support these components of effective WTL, and considerations regarding these components of WTL pedagogy guided the conceptualization of the automated feedback tool presented herein. It is important to study whether and how the automated feedback tool supports these features of effective WTL.

Although the literature demonstrates the benefits of WTL, several barriers can prevent instructors from adopting the practice.[9,10] One notable barrier is the time it takes to provide students with feedback, which can be especially challenging when implementing writing in large-enrollment courses.[9,10] Therefore, it is necessary to research strategies to overcome this barrier. One approach is to use ML to generate formative feedback in order to overcome the time limitation associated with instructors directly providing students feedback. Using ML to analyze writing in the context of STEM instruction is gaining increasing attention in the literature, with the goal of examining student writing to gain insight into students' conceptual understanding and knowledge.[11–17] However, studies using this methodology largely focus on scoring student responses and validating the ML approach by comparing human and machine scores.[11] Hence, it is necessary to further this area of research by contributing to the few existing studies which report developing and implementing tools that can provide automated feedback to support students' learning.[18] Existing research on undergraduate-level chemistry courses demonstrates one approach by automatically scoring responses to short response items and presenting students with adaptive tutorials based on their score.[19,20] The existing research describing automated feedback tools for student writing largely focuses on providing feedback on the quality of students' writing, such as their argumentation or academic writing skills, rather than providing feedback to support students' learning of content knowledge.[21–24] As such, there is a need to research automated evaluation of student writing to provide content-focused, formative feedback. As such, the goal of this contribution is to work towards the development of automated feedback tools which use ML to promote students' content learning during the process of drafting and revising WTL assignments.

In addition to conceptualizing the automated feedback tool with consideration of the effective WTL components, we aligned the development of the feedback tool with the theory of self-regulated learning.[25] This theory proposes that self-regulated processes (such as cognition and behavior) mediate how students use their domain knowledge to set goals and develop strategies towards achieving internal learning outcomes for a given task. When students produce a response to the task, they can receive external, formative feedback (typically from the instructor), which can influence their internal processes. Teaching strategies that can promote self-regulated learning include providing high-quality feedback that clarifies good performance, facilitates self-assessment, and encourages dialogue. Self-regulated learning theory supports WTL pedagogy, particularly when WTL requires students to revise their writing after submitting an initial draft. Furthermore, developing tools for automated feedback can provide a means to enact teaching strategies that can promote students' self-regulated learning; for example, the nature of the feedback can clarify expectations for the assignment and promote self-assessment as students work on revising their writing. As such, this theory guided the development of the feedback tool described in this article, with the goal of developing a tool that can provide feedback that promotes students' self-regulated learning.

10.5 Method

The methodology for the feedback tool design is reported in alignment with the technologies for automated feedback classification framework, which provides guidelines for reporting the educational context, feedback properties, architecture, and evaluation of automated feedback tools.[26]

10.5.1 Educational context

Domain, level, and setting. This research is within the context of a second-semester introductory organic chemistry laboratory course at a large Midwestern research university that typically enrolls 800 students yearly. The course is taught as an in-person lecture by faculty and postdoctoral instructors, with several smaller laboratory sections led by graduate student instructors. The WTL assignment for which the tool was developed is one of three in the course. The WTL assignments require students to submit initial drafts, participate in anonymous and asynchronous peer review, and submit revised drafts. After the assignment is given to students, they have one week to submit their initial drafts. During peer review, students receive feedback from and provide feedback to approximately three peers, based on content-focused evaluation criteria. After peer review, students have three days to revise and submit their final drafts. During the weeks in which the WTL assignments are open, students have access to writing fellows, who are upper-level undergraduate students trained to support students with the WTL assignments. The writing fellows hold office hours and respond to student questions *via* email. The peer review process and writing fellows are in place to support the effective implementation of WTL.[27] The materials associated with the WTL assignment (i.e., the assignment text, peer review criteria, and the evaluation rubric) are all posted on the course's online learning management system.

10.5.2 Feedback properties

Purpose, adaptiveness, timing, and learner control. The automated feedback tool is intended to provide students with content-focused, formative feedback while students work on their responses to the WTL assignment outside of class time. The feedback is intended to be both formative and suggestive: formative because students are intended to use the tool while they still have the opportunity to revise their writing, and suggestive because the feedback provides information about students' written responses in relation to specific criteria (described below) while providing guidance for their revisions. Because of the individualized nature of the feedback generation model (described below), the tool is task adaptive, in that the feedback is adapted to individual students' responses. Students have control over when they receive feedback, as the tool is intended to be available at any time from when students are introduced to the assignment to when they submit their final draft and can be used as often as students wish throughout the WTL process.

10.5.3 Architecture

Student data. The feedback tool is developed for students to use with their written drafts to one of the WTL assignments in the course. The assignment elicited students' descriptions and explanations of two organic chemistry reaction mechanisms, in alignment with a course goal to promote students' understanding of the reactivity of organic compounds. Students using the tool can submit their response, which is automatically analyzed with ML models to provide students with tailored feedback based on their writing. The ML models, which are reported in prior work, identify features of students' writing pertaining to their descriptions and explanations of organic reaction mechanisms; the models were developed using student data from prior administrations of the WTL assignment.[28]

Domain model. The domain model aligns with a framework for identifying evidence of mechanistic reasoning.[29] Prior research elaborates this framework in the context of organic chemistry reaction mechanisms, demonstrating its utility for identifying features of students' writing relevant to reasoning with reaction mechanisms.[4,28,30] In alignment with the goal to provide students with feedback to support their development of more thorough explanations of reaction mechanisms, the domain model guided the specific features identified within students' writing that served as the basis for the automated feedback. Specifically, the domain model identified ten features of mechanistic reasoning (Figure 10.1). Previous research describes the development of the individual ML models for each feature using convolutional neural networks with the Scikit-learn and Keras libraries in Python 3.[28,31,32] The prior research indicates that the ML models for the features included in the feedback tool reached moderate to strong agreement with human raters, with accuracy values, Cohen's kappa values, and Matthews correlation coefficients between 0.884–0.997, 0.738–0.993, and 0.739–0.993, respectively.[28]

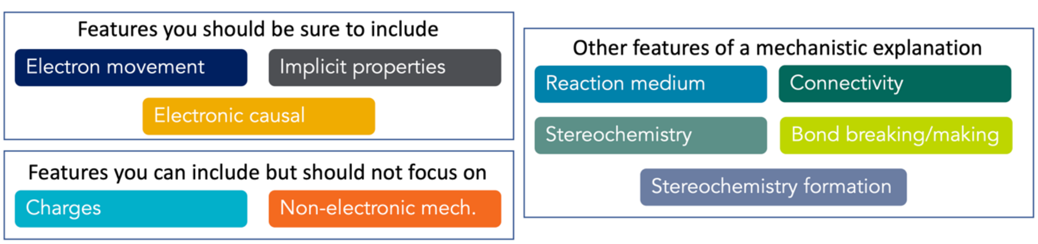


Figure 10.1 The domain model, informed by expert knowledge, for the ML models used to automatically analyze student writing. The abbreviation "mech." is for "mechanism."

Expert knowledge. As chemistry education researchers focused on organic chemistry education, we used our expert knowledge of the domain to develop rules for generating the formative feedback and to articulate specific feedback statements. We first identified the features of the domain model as either (1) features that students should focus on including, which are representative of more sophisticated reasoning; (2) features that students could include but should not focus on, which are representative of surface-level reasoning; or (3) the other features for explaining mechanisms, as a neutral category (Figure 10.1). We then developed feedback statements based on the feedback we would give a student dependent upon the extent to which they elaborated on each feature.

Feedback generation model. To generate the individualized feedback provided to students, the number of sentences identified for each feature in their response is compared to the average and 25th quartile for the number of sentences per feature across the full dataset of students' responses from prior implementations of the assignment. For two features, if a student's response was above average, they were provided with feedback under the header, "For your revisions, consider how you can be more specific in your explanations or reasoning by addressing the following feature(s)." These two features represent aspects of students' mechanistic reasoning reflective of surface-level understanding that could be revised to include deeper-level reasoning. Therefore, feedback statements were provided in this scenario to encourage more specific explanations. For the remaining eight features, if the student response was below the 25th quartile, the feedback was placed under the header, "For your revisions, you are encouraged to focus on including more of the following feature(s)," whereas if their response was below average, but above the 25th quartile, the feedback statement was placed under the header, "For your revisions, you may also want to focus on including more of the following feature(s)." In the scenario where students had no features below average, they were provided with the statement,

> "Compared to the average response, your response includes an above average number of references to each feature identified. Consider revising your response to make sure you are not being too lengthy in your writing. Furthermore, ensure that your mechanistic explanation includes a description of electron movement justified by the chemical/physical properties of the reacting species."

Implementation. The feedback tool was developed in Streamlit, an open-source platform for designing web-based applications in Python.[33] While the tool has not yet been implemented into the course, the tool is designed to eventually link the tool to the course's learning management system.

10.5.4 Evaluation

Evaluation method. After developing the prototype, we evaluated the feedback tool through pilot interviews with students. The data collection procedures were granted exempt status by the Institutional Review Board. At the beginning of the term, we administered a survey to identify students who would be willing to be interviewed for research purposes; we contacted these students after they had completed all portions of the relevant WTL assignment. Of the students contacted, five responded and consented to participate in an interview. The interviewed students were all in their second year of their undergraduate programs, all with majors in biology-related programs. One student was pursuing a second degree in computer science. All students were enrolled in the organic chemistry laboratory course to fulfill requirements for their degree programs. During the interviews, students were provided time to use the feedback tool with their drafts from the assignment, following which we sought their input on the design and usability of the tool and how students would use the tool if it were available to them. All interviews were audio recorded and transcribed verbatim. The transcripts were analyzed through deductive coding[34] based on the components of effective WTL pedagogy[1] and the theory for self-regulated learning.[25] One researcher (FMW) iteratively coded the interview transcripts, and the coding scheme was finalized through discussions with the research team. A second researcher (AJD) then independently read the analyzed transcripts and verified the coding, followed by both researchers discussing the analysis. One researcher (FMW) then thematically analyzed the coded transcripts, identifying and defining themes in alignment with the components of effective WTL pedagogy. The research team engaged in discussions throughout the thematic analysis to ensure consistency in the interpretation of the interview transcripts. Pseudonyms are used when discussing interview responses.

10.6 Results and Discussion

10.6.1 The automated feedback tool

The prototype of the automated feedback tool is shown in Figure 10.2. When students open the tool, they first see text describing the goals of the writing assignment and how to use the feedback tool (Figure 10.2A). Under the instructions is a textbox where students can copy and paste a draft of their writing assignment. Once students click "Submit," the tool analyzes their writing and generates feedback. Students first see graphical feedback in the form of a color-coded bar chart indicating the number of sentences identified within their response pertaining to the necessary features of mechanistic reasoning (Figure 10.2B). Students can open a sidebar on the left side of the application window to view definitions for each feature along the *x*-axis of the graph (Figure 10.2C). Below the graph, students receive tailored written feedback for specific features

(Figure 10.2D). The graphical and written feedback are intended to support self-regulated learning by highlighting the features students should include in their writing to help students identify what is expected in their response.[25] This can facilitate reflection on their writing and support their self-assessment and revisions.[25]

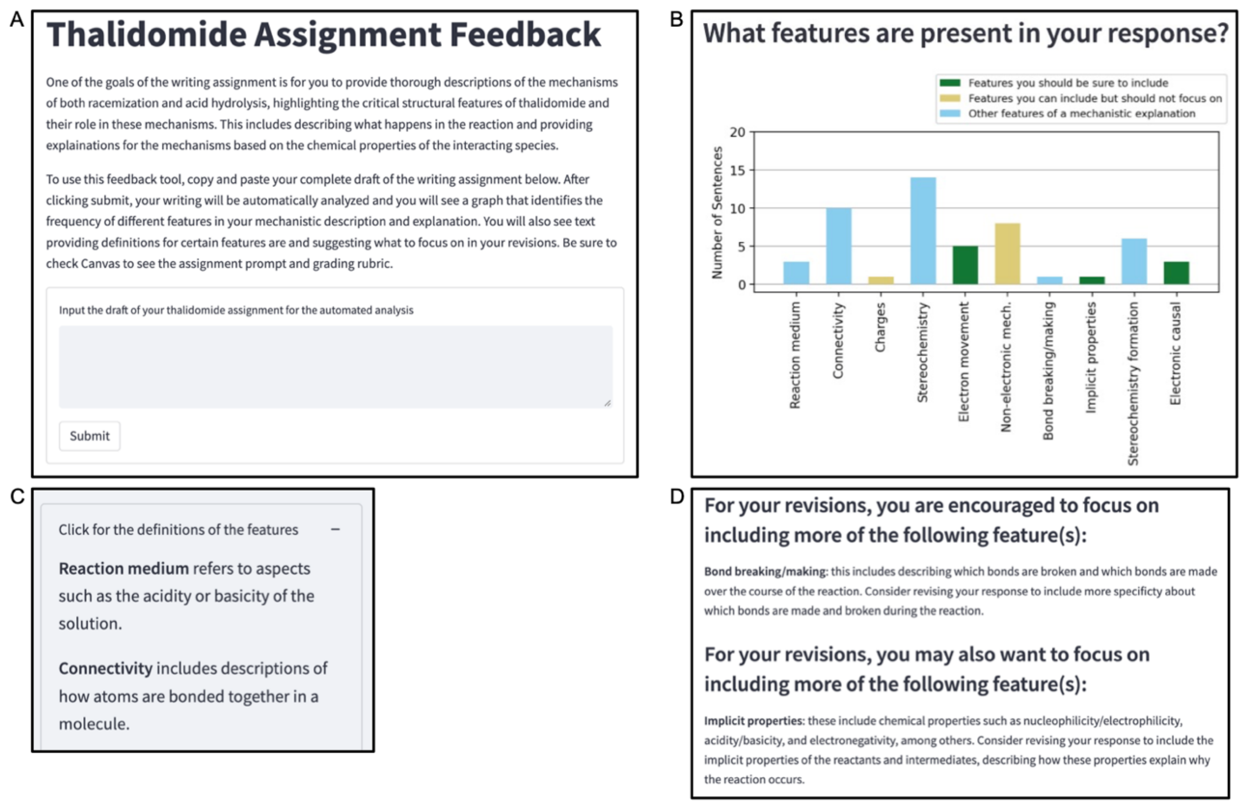


Figure 10.2 Prototype of the automated feedback tool: (A) Instructions for how to use the tool and a text box for students to submit their drafts; (B) the graphical feedback display; (C) an excerpt of the sidebar with definitions of features; (D) the tailored written feedback.

10.6.2 Evaluation of the automated feedback tool

Analysis of the pilot interviews with students using the feedback tool indicated that students' perceptions largely aligned with the effective components of WTL pedagogy and the key aspects of feedback to promote self-regulated learning.[1,25] Specifically, the interviews elicited students' perceptions of how the feedback tool can (1) communicate expectations for their responses to the assignments; (2) promote their metacognitive processes; and (3) influence their interactive writing processes. These three themes highlight how the feedback tool can clarify good performance, facilitate self-assessment, and encourage dialogue, respectively, which are key aspects suggested to promote students' self-regulated learning.[25]

The feedback tool can communicate expectations for students' responses to the assignments. All students indicated finding the tool helpful for signaling the information that should be included in their writing, in alignment with the goal to support self-regulated learning by providing high quality feedback that clarifies expectations for students' performance.[25] For example, Alex stated,

> "It shows you where you can be more specific… it tells you, 'You might want to focus more on this.' And then also in the graph, it shows what you should definitely include… I think that kind of outlines the expectations for… what you should have."

Four students (Taylor, Emerson, Alex, and Jordan) discussed that including the definitions of terms in the sidebar helped clarify the meaning of the feedback. For example, Taylor stated, "The definitions on the site are really great… these would be really good if there's… any type of uncertainty as to what something means." Two students (Emerson and Alex) suggested that it would also be helpful if the tool could identify specific sentences that were tagged with each feature in order to see examples from their own writing where they successfully incorporated each feature. Four students (Taylor, Emerson, Alex, and Jordan) explicitly discussed using the tool in conjunction with the assignment text and/or rubric to clarify the expectations for their responses to the assignment. For instance, Taylor said,

> "I think I would have to use this tool in addition to looking at the rubric. I think the rubric is more specific about [indicating], 'we want to see the mechanisms,' and 'we want to see why it would happen'… but this helps you kind of get to that bigger goal that the rubric is asking for."

These students' comments suggested that the tool could be improved by including tabs to display the assignment text and rubric.

When discussing the graphical feedback specifically, all students indicated that the color-coding indicating the importance of different features was also helpful for guiding their thinking about what should be included in their response. For instance, Alex stated,

> "It was nice just to see the three categories… I like the like color coding, and it's really, like simply set up to see what you have and what you don't have, and able to connect that to like the feedback that it gives you in writing."

However, three students (Taylor, Ryan, and Jordan) brought up questions about the goals for their performance and suggested these goals be better articulated. This concern is encapsulated by Jordan's statement that,

> "When you give someone the ability to run a numerical output, there is a want to make that numerical output hit… a certain shape… I feel like I would try to essentially see how high I can get all of these bars, and at that point, I end up turning in like a ten-page paper."

Two other students, Taylor and Ryan, shared a similar concern about what distribution of features they should aim to incorporate in their writing. Taylor and Ryan indicated that incorporating numerical goals for what students should include in their responses would alleviate this concern. These students' feedback indicates improvements that can better signal expectations for students' performance and support self-regulated learning.[25]

The feedback tool can promote students' metacognitive processes. Four students (Taylor, Emerson, Alex, Jordan) discussed how the feedback would promote reflection on their responses by prompting them to consider whether they included the specific features in their response. For example, Taylor stated,

> "At first when I was looking at it I was like, 'Oh, this is important to have. Did I have it'?... If I was doing like, draft one, and then I saw this, I would probably go back and read it all and look for how many times did I mention charges?"

These students further discussed how reflection prompted by the feedback would guide their thought processes for revising their writing. For instance, Emerson stated, "If I was actually writing this, I could probably go back and be like, Okay, let me talk more about like, electron movement, because I probably barely mentioned it." Similarly, Alex stated,

> "Reading these definitely, like puts it into my mind that like, I need to change how I talk about these certain features or need to add them. And I would, I guess, simply just go through my draft again, kind of highlight stuff that could be changed, or places I need to add, or clarify, and then just kind of work from there."

The indication that students would use the feedback to reflect on their writing and guide their revisions indicates the possible value of the tool for promoting students' self-assessment and engagement in self-regulated learning.[25]

The feedback tool can influence students' interactive writing processes. Four students (Taylor, Ryan, Emerson, and Alex) discussed how the feedback tool would influence their engagement with the structures in place to promote interactive writing processes: peer review and the writing fellows' office hours. The students discussed how they would consider both the peer feedback and the automated feedback together. For example, Ryan stated, "I would probably, like do edits based on peer review first, and then… I would put it in [the automated feedback tool] and then address the stuff that it says to address." Similarly, Taylor said,

> "I would use it before I submit my first draft… I think I would just like screenshot where I was at, and what the suggestions were, and then get my peer review feedback, and then pull up [the automated feedback] again from draft one, and then all together make further changes."

One student, Emerson, noted how they would focus on determining the correctness of their response through peer feedback, while using the automated feedback to consider the thoroughness of their description and explanation of the reaction mechanisms.

There was a divide between students regarding whether the feedback tool would promote their attendance during writing fellows' office hours. Two students (Taylor and Ryan) who indicated rarely attending office hours stated that they would be less likely to attend if the feedback tool were available. In contrast, two other students (Emerson and Alex), both of whom indicated typically attending office hours, stated that they would continue to use the resource alongside the automated feedback. For instance, Alex stated,

> "I would still definitely use the [writing] fellows in their office hours, just to make sure that I'm hitting like the major points in the paper… if I had the [automated] feedback, I would probably like submit a draft first, and see what feedback it has, before I went to the fellows, because I think it might save me time."

Students' responses to how they would still engage in peer review and use writing fellows' office hours suggests that the feedback tool would not interfere with existing structures to promote interactive writing, and would encourage teacher and peer dialogue, further supporting students' self-regulated learning.[1,25]

A final idea that arose during one interview was how the automated feedback tool would provide a means for interactive writing when students might otherwise face challenges going to office hours. Alex stated,

> "I know personally that I have a lot of trouble, like, asking for help and like, or like going to office hours… automatic feedback that it's like, right away, is super helpful, I think for students to like, get a general idea of what they might need to revise in their drafts."

This idea suggests that the automated feedback tool can provide a resource for interactive writing for students who may be unable to attend office hours (for reasons such as anxiety, scheduling conflicts, etc.). While this idea only arose during a single interview, it suggests the possibility of the feedback tool enabling increased access to support for completing the assignment, engaging in the writing process, and developing self-regulated learning skills.[25]

10.7 Conclusion and future work

This work describes the development of a prototype tool for automated, formative feedback to students' responses to a content-specific WTL assignment in organic chemistry. The tool was developed in alignment with the theory of self-regulated learning to complement the effective components of WTL pedagogy. The design of the feedback tool extends the state-of-the-art for writing analytics by demonstrating an approach for automatically delivering content-specific formative feedback on student writing, which is important for supporting self-regulated learning within content-focused pedagogies. Pilot interviews with students using the feedback tool indicate generally positive alignment with the effective components of WTL. Future work will involve further development of specific aspects of the tool based on students' suggestions during the interviews. These improvements will include straightforward functionality improvements such as communicating the categorization of each feature's importance for mechanistic reasoning elsewhere on the bar graph along with indicating target goals for how often students should aim to include each feature in their writing. Further development will also include working on a feature students can use to see specific sentences which were tagged with each feature. After working on these improvements, future research will involve implementing the tool for all students to use during the course and collecting data to better understand how students use the feedback and how to tool interfaces with WTL pedagogy within the classroom setting. This research agenda will contribute to developing and implementing approaches for delivering automated feedback using ML to promote students' content learning in organic chemistry and STEM courses more broadly.

10.8 References

(1) Gere, A. R.; Limlamai, N.; Wilson, E.; MacDougall Saylor, K.; Pugh, R. Writing and

Conceptual Learning in Science: An Analysis of Assignments. *Writ. Commun.* **2019**, *36* (1), 99–135. https://doi.org/10.1177/0741088318804820.

(2) Anderson, P.; Anson, C. M.; Gonyea, R. M.; Paine, C. The Contributions of Writing to Learning and Development: Results from a Large-Scale Multi-Institutional Study. *Res. Teach. English* **2015**, *50* (2), 199–235.

(3) Schmidt-McCormack, J. A.; Judge, J. A.; Spahr, K.; Yang, E.; Pugh, R.; Karlin, A.; Sattar, A.; Thompson, B. C.; Gere, A. R.; Shultz, G. V. Analysis of the Role of a Writing-To-Learn Assignment in Student Understanding of Organic Acid-Base Concepts. *Chem. Educ. Res. Pract.* **2019**, *20* (2), 383–398. https://doi.org/10.1039/c8rp00260f.

(4) Watts, F. M.; Schmidt-McCormack, J.; Wilhelm, C.; Karlin, A.; Sattar, A.; Thompson, B.; Gere, A. R.; Shultz, G. What Students Write about When Students Write about Mechanisms: Analysis of Features Present in Students' Written Descriptions of an Organic Reaction Mechanism. *Chem. Educ. Res. Pract.* **2020**, *21*, 1148–1172. https://doi.org/10.1039/C9RP00185A.

(5) Watts, F. M.; Park, G. Y.; Petterson, M. N.; Shultz, G. V. Considering Alternative Reaction Mechanisms: Students' Use of Multiple Representations to Reason about Mechanisms for a Writing-to-Learn Assignment. *Chem. Educ. Res. Pract.* **2022**, *23*, 486–507. https://doi.org/10.1039/d1rp00301a.

(6) Brandfonbrener, P. B.; Watts, F. M.; Shultz, G. V. Organic Chemistry Students' Written Descriptions and Explanations of Resonance and Its Influence on Reactivity. *J. Chem. Educ.* **2021**. https://doi.org/10.1021/acs.jchemed.1c00660.

(7) Gupte, T.; Watts, F. M.; Schmidt-McCormack, J. A.; Zaimi, I.; Gere, A. R.; Shultz, G. V. Students' Meaningful Learning Experiences from Participating in Organic Chemistry Writing-to-Learn Activities. *Chem. Educ. Res. Pract.* **2021**, *22*, 396–414. https://doi.org/10.1039/d0rp00266f.

(8) Petterson, M. N.; Finkenstaedt-Quinn, S. A.; Gere, A. R.; Shultz, G. V. The Role of Authentic Contexts and Social Elements in Supporting Organic Chemistry Students' Interactions with Writing-to-Learn Assignments. *Chem. Educ. Res. Pract.* **2022**. https://doi.org/10.1039/d1rp00181g.

(9) Finkenstaedt-Quinn, S. A.; Gere, A. R.; Dowd, J. E.; Thompson Jr., R. J.; Halim, A. S.; Reynolds, J. A.; Schiff, L. A.; Flash, P.; Shultz, G. V. Postsecondary Faculty Attitudes and Beliefs About Writing-Based Pedagogies in the STEM Classroom. *CBE Life Sci. Educ.* **2022**, Accepted.

(10) Thompson, R. J.; Finkenstaedt-Quinn, S. A.; Shultz, G. V.; Gere, A. R.; Schmid, L.; Dowd, J. E.; Mburi, M.; Schiff, L. A.; Flash, P.; Reynolds, J. A. How Faculty Discipline and Beliefs Influence Instructional Uses of Writing in STEM Undergraduate Courses at Research-Intensive Universities. *J. Writ. Res.* **2021**, *12* (3), 625–656. https://doi.org/10.17239/jowr-

2021.12.03.04.

(11) Zhai, X.; Yin, Y.; Pellegrino, J. W.; Haudek, K. C.; Shi, L. Applying Machine Learning in Science Assessment: A Systematic Review. *Stud. Sci. Educ.* **2020**, *56* (1), 111–151. https://doi.org/10.1080/03057267.2020.1735757.

(12) Zhai, X. Practices and Theories: How Can Machine Learning Assist in Innovative Assessment Practices in Science Education. *J. Sci. Educ. Technol.* **2021**, *30* (2), 139–149. https://doi.org/10.1007/s10956-021-09901-8.

(13) Ha, M.; Nehm, R. H.; Urban-Lurain, M.; Merrill, J. E. Applying Computerized-Scoring Models of Written Biological Explanations across Courses and Colleges: Prospects and Limitations. *CBE Life Sci. Educ.* **2011**, *10* (4), 379–393. https://doi.org/10.1187/cbe.11-08-0081.

(14) Haudek, K. C.; Kaplan, J. J.; Knight, J.; Long, T.; Merrill, J.; Munn, A.; Nehm, R.; Smith, M.; Urban-Lurain, M. Harnessing Technology to Improve Formative Assessment of Student Conceptions in STEM: Forging a National Network. *CBE Life Sci. Educ.* **2011**, *10* (2), 149–155. https://doi.org/10.1187/cbe.11-03-0019.

(15) Dood, A. J.; Dood, J. C.; Cruz-Ramírez De Arellano, D.; Fields, K. B.; Raker, J. R. Analyzing Explanations of Substitution Reactions Using Lexical Analysis and Logistic Regression Techniques. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 267–286. https://doi.org/10.1039/c9rp00148d.

(16) Dood, A. J.; Fields, K. B.; Raker, J. R. Using Lexical Analysis to Predict Lewis Acid-Base Model Use in Responses to an Acid-Base Proton-Transfer Reaction. *J. Chem. Educ.* **2018**, *95* (8), 1267–1275. https://doi.org/10.1021/acs.jchemed.8b00177.

(17) Yik, B. J.; Dood, A. J.; Cruz-Ramírez De Arellano, D.; Fields, K. B.; Raker, J. R. Development of a Machine Learning-Based Tool to Evaluate Correct Lewis Acid-Base Model Use in Written Responses to Open-Ended Formative Assessment Items. *Chem. Educ. Res. Pract.* **2021**, *22* (4), 866–885. https://doi.org/10.1039/d1rp00111f.

(18) Xiong, Y.; Wu, Y.-F. B. An Automated Feedback System to Support Student Learning in Writing-to-Learn Activities. *Proc. 6th 2019 ACM Conf. Learn. Scale, L@S 2019* **2019**, 1–4. https://doi.org/10.1145/3330430.3333658.

(19) Dood, A. J.; Fields, K. B.; de Arellano, D. C. R.; Raker, J. R. Development and Evaluation of a Lewis Acid–Base Tutorial for Use in Postsecondary Organic Chemistry Courses. *Can. J. Chem.* **2019**, *97* (10), 711–721. https://doi.org/10.1139/cjc-2018-0479.

(20) Dood, A. J.; Dood, J. C.; Cruz-Ramírez De Arellano, D.; Fields, K. B.; Raker, J. R. Using the Research Literature to Develop an Adaptive Intervention to Improve Student Explanations of an SN1 Reaction Mechanism. *J. Chem. Educ.* **2020**, *97* (10), 3551–3562. https://doi.org/10.1021/acs.jchemed.0c00569.

(21) Larrondo, P.; Frank, B.; Ortiz, J. The State of the Art in Providing Automated Feedback To Open-Ended Student Work. *Proc. Can. Eng. Educ. Assoc.* **2021**. https://doi.org/10.24908/pceea.vi0.14854.

(22) Ngo, T. T.-N.; Chen, H. H.-J.; Lai, K. K.-W. The Effectiveness of Automated Writing Evaluation in EFL/ESL Writing: A Three-Level Meta-Analysis Writing. *Interact. Learn. Environ.* **2022**, 1–18. https://doi.org/10.1080/10494820.2022.2096642.

(23) Knight, S.; Shibani, A.; Abel, S.; Gibson, A.; Ryan, P.; Sutton, N.; Wight, R.; Lucas, C.; Sándor, Á.; Kitto, K.; et al. AcaWriter: A Learning Analytics Tool for Formative Feedback on Academic Writing. *J. Writ. Res.* **2020**, *12* (1), 0–2.

(24) Strobl, C.; Ailhaud, E.; Benetos, K.; Devitt, A.; Kruse, O.; Proske, A.; Rapp, C. Digital Support for Academic Writing: A Review of Technologies and Pedagogies. *Comput. Educ.* **2019**, *131*, 33–48. https://doi.org/10.1016/j.compedu.2018.12.005.

(25) Nicol, D.; MacFarlane-Dick, D. Formative Assessment and Selfregulated Learning: A Model and Seven Principles of Good Feedback Practice. *Stud. High. Educ.* **2006**, *31* (2), 199–218. https://doi.org/10.1080/03075070600572090.

(26) Deeva, G.; Bogdanova, D.; Serral, E.; Snoeck, M.; De Weerdt, J. A Review of Automated Feedback Systems for Learners: Classification Framework, Challenges and Opportunities. *Comput. Educ.* **2021**, *162* (March 2020), 104094. https://doi.org/10.1016/j.compedu.2020.104094.

(27) Finkenstaedt-Quinn, S. A.; Petterson, M.; Gere, A.; Shultz, G. Praxis of Writing-to-Learn: A Model for the Design and Propagation of Writing-to-Learn in STEM. *J. Chem. Educ.* **2021**, *98* (5), 1548–1555. https://doi.org/10.1021/acs.jchemed.0c01482.

(28) Watts, F. M.; Dood, A. J.; Shultz, G. V. Exploring Students' Written Descriptions of Organic Reaction Mechanisms and Developing an Automated Approach for Analysing Students' Writing. In *Student Reasoning in Organic Chemistry*; Graulich, N., Shultz, G. V., Eds.; Royal Society of Chemistry, 2022.

(29) Russ, R. S.; Scherr, R. E.; Hammer, D.; Mikeska, J. Recognizing Mechanistic Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed from Philosophy of Science. *Sci. Educ.* **2008**, *92* (3), 499–525. https://doi.org/10.1002/sce.20264.

(30) Dood, A. J.; Watts, F. M. Mechanistic Reasoning in Organic Chemistry: A Scoping Review of How Students Describe and Explain Mechanisms in the Chemistry Education Research Literature. *J. Chem. Educ.* **2022**, *99* (8), 2864–2876. https://doi.org/10.1021/acs.jchemed.2c00313.

(31) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in

Python Fabian. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(32) Chollet, F.; others. Keras. 2015.

(33) Treuille, A.; Teixeira, T.; Kelly, A. Streamlit.

(34) Miles, M. B.; Huberman, A. M.; Saldana, J. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed.; Sage: Los Angeles, CA, 2014.

Chapter 11

Closing Remarks

The body of research presented within this dissertation represents significant contributions to the chemistry education research literature focused on eliciting and supporting students' reasoning in organic chemistry. The two literature review chapters (Chapters 1 and 4) synthesize the literature on mechanistic reasoning and the MWrite writing-to-learn (WTL) research, respectively, providing insight into the current state of knowledge with respect to these two major research areas of the dissertation. The studies presented at the beginning of the dissertation (Chapters 2 and 3) provide insight into how students reason with specific organic chemistry reaction mechanisms while also demonstrating the utility of different instructional tools to support and elicit students' reasoning. The research focused on WTL, which make up a majority of the dissertation (Chapters 5 through 8), constitutes a major portion of the research on WTL assignments within the organic chemistry context. This collection of studies specifically demonstrates the value of WTL for both providing students with meaningful learning experiences while also promoting their conceptual engagement and development of reasoning skills. The final section of the dissertation (Chapters 9 and 10), focused on developing machine learning models to provide students with automated, formative feedback on the WTL assignments, contributes to a new and growing area of research within the field of chemistry education focused on leveraging machine learning. These studies represent some of the initial work within the field demonstrating how machine learning technology can support the evaluation and implementation of formative assessments, such as WTL, in chemistry courses. The findings across the collection of studies in the dissertation highlight the various nuances in students' mechanistic reasoning and conceptual understanding, demonstrate the value of eliciting students' understanding, and present strategies for leveraging students' elicited reasoning to better support their learning.